

Перв. примен.	Оглавление				
	<div>Введение 4</div> <div>Глава 1. Обзор существующих методов..... 7</div> <div>1.1. Структура алгоритмов локализации объектов на изображении 7</div> <div>1.2. Обзор существующих методов локализации объекта на изображении.. 9</div> <div>1.2.1. Методы построения признакового описания изображений 10</div> <div>1.2.2. Методы локализации объектов на изображении 16</div> <div>1.3. Выводы по главе 1 22</div> <div>Глава 2. Информационная модель локализации объектов на изображении 24</div> <div>2.1. Теория активного восприятия 24</div> <div>2.1.1. U-преобразование 24</div> <div>2.1.2. Алгебра групп..... 29</div> <div>2.2. Реализация системы локализации объекта на изображении на основе теории активного восприятия..... 32</div> <div>2.2.1. Предварительная обработка сигнала 32</div> <div>2.2.2. Формирование признакового описания искомого объекта 33</div> <div>2.2.3. Принятие решения о локализации объекта на изображении 36</div> <div>2.3. Выводы по главе 2 42</div> <div>Глава 3. Вычислительный эксперимент 43</div> <div>3.1. Описание эксперимента 43</div> <div>3.2. Описание программного продукта 46</div> <div>3.3. Анализ результатов эксперимента 50</div> <div>3.4. Выводы по главе 3 54</div> <div>Заключение..... 55</div> <div>Список литературы 56</div>				
Справ. №					
Подпись и дата					
Инв. № дубл.					
Взам. инв. №					
Подпись и дата					
Инв. № подл.					

					ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)						
Изм.	Лист	№ докум.	Подпись	Дата							
Разраб.		Филяков А.А.			Модель и алгоритмы локализации объектов на изображении						
Провер.		Гай В.Е.									
Реценз.											
Н. Контр.											
Утверд.											
					<table><tr><td>Лит.</td><td>Лист</td><td>Листов</td></tr><tr><td></td><td>3</td><td>57</td></tr></table>	Лит.	Лист	Листов		3	57
Лит.	Лист	Листов									
	3	57									
					НГТУ им. Р.Е. Алексеева						

Перв. примен.	Введение														
Справ. №	<p>Актуальность темы</p> <p>Задача локализации объектов на изображении является одной из распространенных задач теоретической информатики. Данная задача имеет широкое применение в областях компьютерного зрения и обработки изображений. Среди таких применений можно выделить распознавание лиц, людей, автомобилей, сооружений и других объектов на фотографиях. Решение подобных задач позволяет формировать машинным способом большие базы данных изображений с подробным описанием их содержимого для последующего поиска этих изображений.</p> <p>В настоящее время существует множество методов, способных решать задачу локализации объекта на изображении. Как правило, этим методам приходится иметь дело с большим объемом входных данных, т.е. исследуемых изображений. Таким образом, эффективность работы методов решения рассматриваемой задачи сильно зависит от выбранного способа представления изображения, т.е. от его признакового описания. Именно используемое признаковое описание изображения и является проблемным местом многих методов решения данной задачи.</p> <p>Исходя из этого, можно сделать вывод о том, что разработка новых моделей и методов для решения задачи локализации объектов на изображении является вполне актуальной.</p>														
Подпись и дата	<p>Цель работы и задачи исследования</p> <p>Целью данной работы является разработка и исследование новых моделей и методов решения задачи локализации объекта на изображении. Для достижения этой цели необходимо решить следующие задачи:</p> <ul style="list-style-type: none"> • Рассмотрение и анализ существующих известных методов решения данной задачи. • Создание информационной модели описания объекта и его локализации на изображении. • Создание алгоритма локализации объекта на изображении. • Проведение вычислительного эксперимента для установления корректности работы созданных моделей и алгоритмов. 														
Подпись и дата	<p>Объект исследования</p> <p>Объектом исследования являются изображения, представленные в одном из доступных цифровых форматов.</p>														
Инв. № подл.	<table border="1"> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>Изм.</td> <td>Лист</td> <td>№ докум.</td> <td>Подпись</td> <td>Дата</td> </tr> </table>										Изм.	Лист	№ докум.	Подпись	Дата
Изм.	Лист	№ докум.	Подпись	Дата											

Перв. примен.	<p>Предмет исследования</p> <p>Предметом исследования являются модели и методы решения задачи локализации объектов на изображении.</p>				
	<p>Методы исследования</p> <p>В ходе выполнения работы были использованы методы формирования признаков описания изображений на основе теории активного восприятия, методы принятия решения о локализации объекта на изображении на основе алгоритмов кластеризации точек. Также был использован метод вычислительного эксперимента, для проведения которого был разработан программный продукт на языке программирования R.</p>				
Справ. №	<p>Положения, выносимые на защиту</p> <p>На защиту выносятся следующие положения:</p> <ul style="list-style-type: none"> Информационная модель признаков описания изображения искомого объекта, построенная на основе положений теории активного восприятия. Алгоритм локализации объекта на изображении на основе кластеризации точек признаков описания объекта. Результаты вычислительного эксперимента. 				
	<p>Научная новизна</p> <p>Научная новизна предлагаемого метода решения задачи локализации объекта на изображении заключается в использовании новой комбинации подходов к решению различных этапов данной задачи. На этапе формирования признаков описания таким подходом является использование инструментария теории активного восприятия. На этапе принятия решения о локализации объекта на изображении – использование методов кластеризации точек.</p>				
Взам. инв. №	<p>Теоретическая и практическая ценность</p> <p>Теоретическая ценность работы заключается в разработанной информационной модели признаков описания объекта, а также в предложенном алгоритме его локализации на изображении.</p> <p>Практическая ценность работы заключается в созданном программном продукте, реализующем указанный выше алгоритм, а также в результатах проведенного вычислительного эксперимента.</p>				
	<p>Изм. Лист № докум. Подпись Дата</p>				
Инв. № подл.	<p>ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)</p>				
	<p>Лист 5</p>				

Перв. примен.	<h3>Апробация работы</h3> <p>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.</p>				
	Справ. №	<h3>Обоснованность и достоверность</h3> <p>Обоснованность и достоверность проведенного исследования доказывается проведением вычислительного эксперимента, анализом его результатов, и сравнением их с результатами существующих методов решения данной задачи.</p>			
		<h3>Публикации</h3> <p>Планируется публикация основных положений и результатов данного исследования в виде доклада или тезисов на одной из научных конференций соответствующего профиля.</p>			
		<h3>Объём и структура</h3> <p>Выпускная квалификационная работа состоит из введения, трёх глав основной части, заключения и списка литературы. Общий объём работы – 57 страниц. Диссертация содержит 19 рисунков, 10 таблиц, 28 формул. Список литературы включает в себя 21 наименование.</p>			
Подпись и дата		Инов. № дубл.	Взам. инв. №	Подпись и дата	Инов. № подл.

Перв. примен.	<h2 style="text-align: center;">Глава 1. Обзор существующих методов</h2> <h3 style="text-align: center;">1.1. Структура алгоритмов локализации объектов на изображении</h3> <p>Большинство современных методов, решающих задачу локализации объектов на изображении, включают в себя три основных этапа:</p> <ul style="list-style-type: none"> • Предварительная обработка изображения. • Построение признакового описания искомого объекта. • Локализация объекта на изображении на основе полученного признакового описания. <p>Предварительная обработка изображения является первым этапом в решении данной задачи. Цель этого этапа – представление изображения искомого объекта в виде, подходящем для последующего формирования его признакового описания.</p> <p>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 локализации данного объекта.</p> <p>Важной частью этапа построения признакового описания искомого объекта является поиск его ключевых точек. Выявление ключевых точек на изображении искомого объекта позволяет представить искомый объект в виде набора особых точек, которые являются характерными для данного объекта и могут быть использованы для его локализации в дальнейшем. Выявление особых точек происходит на основе их признакового описания. Данные точки должны быть устойчивы к модификациям масштаба и поворота изображения. Также представление объекта в виде набора ключевых точек позволяет сократить размер его описания, что в свою очередь приводит к снижению объема вычислений на следующем этапе.</p> <p>Третий этап заключается в принятии решения о локализации искомого объекта в месте скопления обнаруженных ключевых точек на исследуемом изображении. Результат данного этапа зависит от выбранного способа принятия такого решения.</p> <p>Схема описанных выше этапов решения задачи представлена на рисунке 1.1:</p>																		
	Справ. №																		
Подпись и дата		Инв. № дубл.	Взам. инв. №	Подпись и дата	Инв. № подл.														
<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 50%; text-align: center; vertical-align: middle;"> ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ) </td> <td style="width: 10%; text-align: center; vertical-align: middle;"> Лист 7 </td> </tr> <tr> <td>Изм.</td> <td>Лист</td> <td>№ докум.</td> <td>Подпись</td> <td>Дата</td> <td colspan="2"></td> </tr> </table>											ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)	Лист 7	Изм.	Лист	№ докум.	Подпись	Дата		
					ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)	Лист 7													
Изм.	Лист	№ докум.	Подпись	Дата															

1.2. Обзор существующих методов локализации объекта на изображении

Построение признакового описания объекта является наиболее важным этапом при решении задачи локализации объекта на изображении. Именно от признакового описания, сложности его формирования и его итогового формата зависит сложность необходимых вычислений на этапе локализации объекта.

Алгоритмы построения признакового описания изображений и поиска их ключевых точек чаще всего заключаются в выявлении на изображении точек, яркость которых является наибольшей или наименьшей по сравнению с соседними точками (светлые пятна на темном фоне или темные пятна на светлом фоне). Далее каждой из таких точек ставится в соответствие описание, которое учитывает величину и направление изменения яркости в данной точке. Среди алгоритмов, формирующих такие описания можно выделить следующие:

- Инвариантное к масштабу признаковое преобразование (SIFT) [1].
- Ускоренное построение устойчивых признаков (SURF) [2][4].
- Двоичные устойчивые независимые элементарные признаки (BRIEF) [3].

Данные алгоритмы отличаются лишь деталями реализации поиска ключевых точек и способом представления признакового описания. Как правило, признаковое описание представляет собой вектор чисел определенной размерности. Главный недостаток таких алгоритмов – значительный объем вычислений, который требуется для формирования признакового описания и его последующей обработки на этапе локализации объекта. Преимуществом является инвариантность получаемых описаний к изменениям масштаба и поворота.

Существует множество различных методов, применяемых для локализации объекта на изображении по ключевым точкам. Среди таких методов можно выделить следующие:

- Локализация объекта на основе объединения областей изображения по принципу схожести [1].
- Иерархическая пирамидальная схема локализации объекта [4].
- Классификатор в виде вероятностных деревьев [6].
- Классификатор k ближайших соседей [7].
- Наивный байесовский классификатор [5].
- Классификатор на основе нейронных сетей [8].

Большинство указанных методов используют предварительное обучение для последующей локализации объекта на изображении. Также некоторые из

Перв. примен.								
Справ. №								
Подпись и дата								
Инв. № дубл.								
Взам. инв. №								
Подпись и дата								
Инв. № подл.								
		Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)	Лист
								9

Перв. примен.					
Справ. №					
Подпись и дата					
Инв. № дубл.					
Взам. инв. №					
Подпись и дата					
Инв. № подл.					

этих методов могут использоваться для классификации объектов по категориям.

1.2.1. Методы построения признакового описания изображений

Данный раздел посвящен методам, использующимся на этапе построения признакового описания изображения. Были подробно рассмотрены следующие алгоритмы:

- SIFT [1].
- SURF [2][4].
- BRIEF [3].

Также указанные алгоритмы решают задачу поиска ключевых точек объекта на изображении.

Алгоритм SIFT

Алгоритм SIFT (Scale invariant feature transform) [1] был представлен в 1999 году как алгоритм, позволяющий выявлять и описывать особые точки на изображениях. Признаковое описание особых точек, формируемое данным алгоритмом, является инвариантным к изменениям масштаба и вращения.

Для выявления и описания ключевых точек, устойчивых к изменениям масштаба используется экстремум функции Гаусса ($D(x, y, \sigma)$), т.е. разница между значениями функции Гаусса ($G(x, y, \sigma)$) при двух соседних масштабах, отличающихся на величину K :

$$D(x, y, \sigma) = (G(x, y, K\sigma) - G(x, y, \sigma)) \times I(x, y) = L(x, y, K\sigma) - L(x, y, \sigma) \quad (1.1)$$

где $D(x, y, \sigma)$ – экстремум функции Гаусса; $G(x, y, \sigma)$ – функция Гаусса; $I(x, y)$ – исследуемое изображение; $L(x, y, \sigma)$ – функция свертки исследуемого изображения с функцией Гаусса.

Процесс, описываемый формулой 1.1 повторяется для участков изображения 8х8 пикселей. Визуализация данного процесса представлена на рисунке 1.2:

					ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)	Лист
						10
Изм.	Лист	№ докум.	Подпись	Дата		

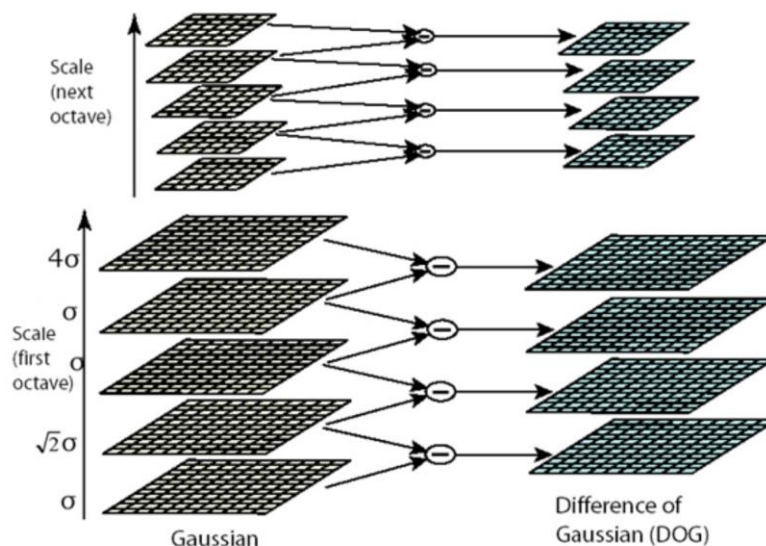


Рисунок 1.2. Визуализация процесса формирования экстремума функции Гаусса $D(x, y, \sigma)$

После этого каждый пиксель сравнивается с 8 соседними пикселями и 18 соседними пикселями из изображений, полученных 2 изменениями масштаба. Если рассматриваемый пиксель имеет наибольшее или наименьшее значение яркости по сравнению со своими соседями, то он является кандидатом на особую точку. Это сравнение не является сложным с точки зрения затрат вычислительной мощности, так как большинство точек отбрасывается после первых нескольких проверок. Данный процесс представлен на рисунке 1.3:

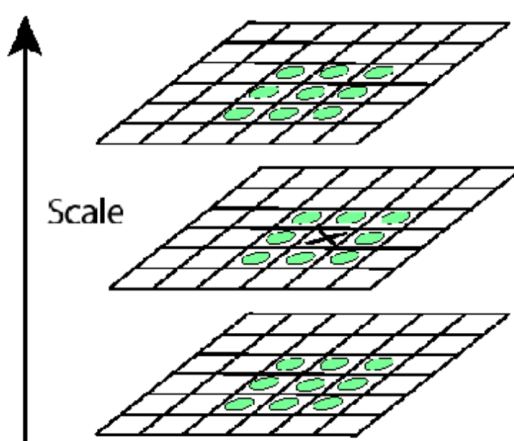


Рисунок 1.3. Процесс сравнения пикселя с соседними для выявления кандидатов на ключевые точки

Перв. примен.					
Справ. №					
Подпись и дата					
Инв. № дубл.					
Взам. инв. №					
Подпись и дата					
Инв. № подл.					

Для каждого кандидата на ключевую точку используется интерполяция соседних данных, чтобы отбросить нестабильные точки, например, точки с низким контрастом.

На следующем шаге каждой ключевой точке присваивается описание ориентации, чтобы получить точки, устойчивые к изменениям вращения. Для этого для каждой точки вычисляется амплитуда (m) и ориентация (θ):

$$m(x, y) = \sqrt{(L(x + 1, y) - L(x - 1, y))^2 + (L(x, y + 1) - L(x, y - 1))^2} \quad (1.2)$$

$$\theta(x, y) = \tan^{-1} \frac{L(x, y + 1) - L(x, y - 1)}{L(x + 1, y) - L(x - 1, y)} \quad (1.3)$$

Пример визуализации полученных локальных градиентов в найденных ключевых точках представлен на рисунке 1.4:

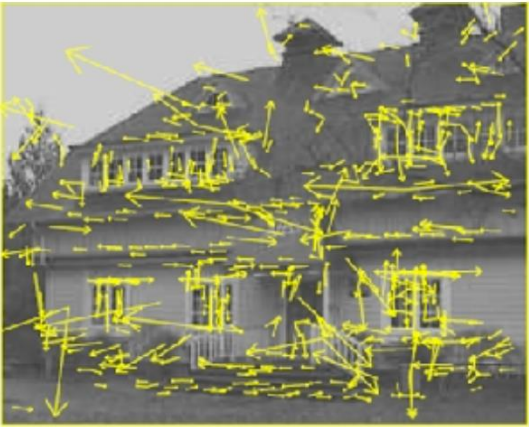


Рисунок 1.4. Локальные градиенты в обнаруженных на изображении ключевых точках

Итоговым признаковым описанием каждой особой точки является вектор размерности 128 (4x4x8), который формируется путем выделения вокруг особой точки 4x4 регионов в 8 направлениях вокруг особой точки. Значение каждого элемента вектора – амплитуда в точке соответствующего региона.

Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)	Лист
						12

Перв. примен.	<h2 style="text-align: center;">Алгоритм SURF</h2> <p>Алгоритм Speeded up robust features (SURF) [2] позволяет выполнять поиск ключевых точек на изображении и строить их признаковое описание. Данный алгоритм основывается на тех же принципах, что и алгоритм SIFT, рассмотренный в предыдущем разделе. Однако стандартная версия SURF работает в несколько раз быстрее, чем SIFT, о чем и говорит его название. Также авторы заверяют, что признаковое описание, получаемое в результате выполнения данного алгоритма, является более устойчивым к различным преобразованиям изображения.</p> <p>SURF использует фильтры квадратной формы как аппроксимацию сглаживания Гаусса в отличие от SIFT, который использует каскадные фильтры. Наиболее быстрая обработка изображения с помощью квадратных фильтров достигается за счет использования суммы изображения:</p> $S(x, y) = \sum_{i=0}^x \sum_{j=0}^y I(i, j) \quad (1.4)$ <p>Поиск особых точек на изображении выполняется на основании матрицы Гессе. Данная матрица вычисляется для каждой точки $p = (x, y)$:</p> $H(p, \sigma) = \begin{pmatrix} L_{xx}(p, \sigma) & L_{xy}(p, \sigma) \\ L_{yx}(p, \sigma) & L_{yy}(p, \sigma) \end{pmatrix} \quad (1.5)$ <p>где $L_{xx}(p, \sigma)$ - результат свертки Гауссова ядра с изображением $I(x, y)$ в точке x.</p> <p>Определитель этой матрицы используется как мера локальных изменений вокруг исследуемой точки. В качестве ключевых выбираются точки, для которых определитель матрицы Гессе принимает большие значения.</p> <p>Важной задачей алгоритма является поиск ключевых точек, устойчивых к изменениям масштаба изображения. Другие алгоритмы часто решают эту задачу с помощью построения пирамиды изображения (SIFT). SURF использует для этого наложение на изображение фильтров с увеличивающимся размером (9x9, 15x15, 21x21, 27x27...).</p> <p>Найденные ключевые точки представляют собой точки с максимальным изменением градиента яркости, т.е. светлые пятна на темном фоне или темные пятна на светлом фоне. Для каждой точки также вычисляется ее</p>																									
	Справ. №																									
Подпись и дата		Инв. № дубл.	Взам. инв. №	Подпись и дата	Инв. № подл.																					
<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 15%;"></td> <td style="width: 15%;"></td> <td style="width: 15%;"></td> <td style="width: 15%;"></td> <td style="width: 15%;"></td> <td style="width: 15%;"></td> <td rowspan="3" style="width: 30%; text-align: center; vertical-align: middle;"> ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ) </td> <td style="width: 10%; text-align: center;">Лист</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td style="text-align: center;">13</td> </tr> <tr> <td style="text-align: center;">Изм.</td> <td style="text-align: center;">Лист</td> <td style="text-align: center;">№ докум.</td> <td style="text-align: center;">Подпись</td> <td style="text-align: center;">Дата</td> <td></td> </tr> </table>												ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)	Лист							13	Изм.	Лист	№ докум.	Подпись	Дата	
						ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)	Лист																			
							13																			
Изм.	Лист	№ докум.	Подпись	Дата																						

Перв. примен.					
Справ. №					
Подпись и дата					
Инв. № дубл.					
Взам. инв. №					
Подпись и дата					
Инв. № подл.					

ориентация – направление наибольшего изменения яркости. Для этого используется фильтр Хаара.

Признаковое описание (дескриптор) каждой ключевой точки строится следующим образом:

- Вокруг точки выбирается окрестность размером $20s$, где s – масштаб изображения, при котором было получено максимальное значение определителя матрицы Гессе, т.е. найден максимальный перепад яркости.
- Данная область разбивается на 4×4 регионов.
- Для каждого региона вычисляется вектор, содержащий 4 элемента: 2 – суммарный градиент по региону, 2 – сумма модулей точечных градиентов.
- Вектор признакового описания особой точки складывается из дескрипторов всех 16 регионов вокруг особой точки. Для большей устойчивости к шумам все элементы вектора взвешиваются на коэффициенты Гауссова ядра.
- К вектору признакового описания добавляется след матрицы Гессе. Это необходимо для выявления светлых и темных пятен, так как след матрицы Гессе имеет различные знаки в случаях светлых и темных пятен.

Таким образом, алгоритм SURF формирует признаковое описание ключевых точек, которое представляет собой вектор, состоящий из 64 элементов. Визуализация данного признакового описания представлена на рисунке 1.5:

Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)	Лист
						14

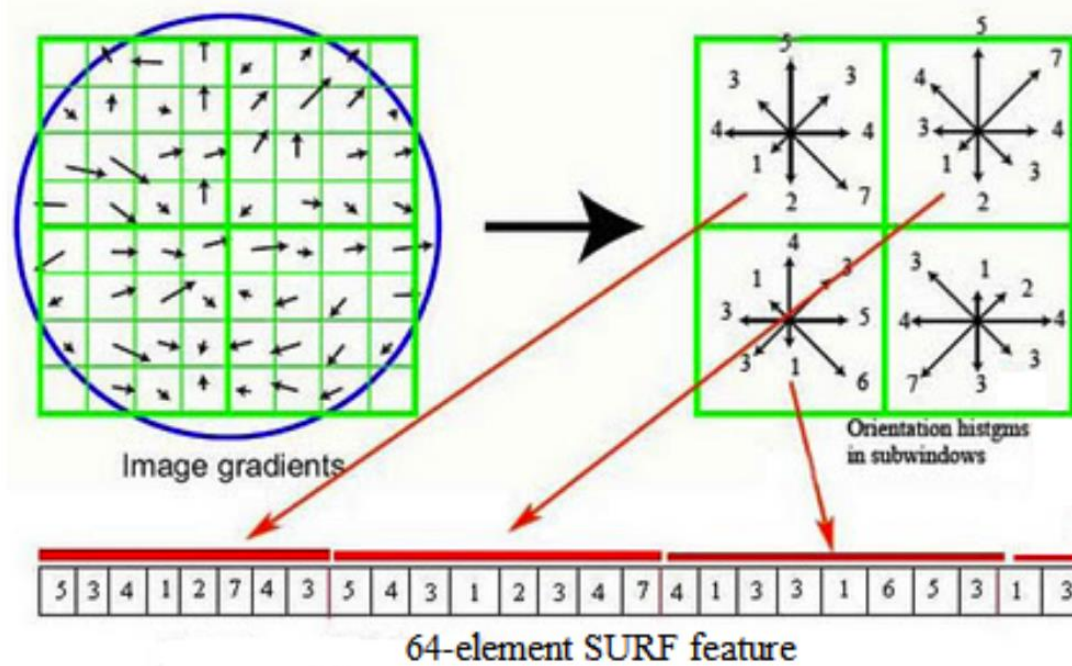


Рисунок 1.5. Визуализация признакового описания точки, полученного методом SURF

Полученное признаковое описание схоже с описанием, получаемым в результате использования алгоритма SIFT, но имеет в 2 раза меньшую размерность. Это позволяет производить вычисления с особыми точками, описанными таким образом, значительно быстрее, что и является главным преимуществом алгоритма SURF, заложенным в его названии.

Алгоритм BRIEF

Алгоритм Binary robust independent elementary features (BRIEF) [3] был предложен как алгоритм, позволяющий выполнять поиск ключевых точек на изображении и формировать их признаковое описание. Главной целью создания данного алгоритма было упрощение формата признакового описания для уменьшения числа дальнейших вычислений. Поэтому в качестве признакового описания были выбраны бинарные строки.

В ходе выполнения данного алгоритма изображение разбивается на перекрывающиеся друг друга участки, называемые областями. Для каждой области p размера $S \times S$ вычисляется так называемый тест t :

Перв. примен.	$\tau(p; x, y) = \begin{cases} 1, & \text{если } p(x) < p(y) \\ 0, & \text{в остальных случаях} \end{cases} \quad (1.6)$				
	<p>где $p(x)$ – интенсивность пикселя x.</p> <p>Таким образом для каждой области формируется набор бинарных тестов, который уникально определяется множеством пар точек (x, y) размерностью n_d. На основании этого набора бинарных тестов вычисляется признаковое описание области, которое представляет собой строку битов размерностью n_d:</p>				
Справ. №	$f_{n_d}(p) = \sum_{1 \leq i \leq n_d} 2^{i-1} \tau(p; x_i, y_i) \quad (1.7)$				
	<p>Величина n_d выбирается в зависимости от требуемой скорости работы алгоритма, используемой памяти и точности распознавания. Наилучший компромисс между этими величинами может быть достигнут при использовании значений n_d 128, 256 или 512. Исходя из этого были получены различные версии алгоритма BRIEF, которые обозначаются как BRIEF- k, где $k = n_d / 8$ – количество байтов, необходимое для хранения одного признакового описания.</p> <p>В качестве меры схожести признаковых описателей, представленных в виде бинарных строк, используется расстояние Хэмминга. Расстояние Хэмминга $d(x, y)$ представляет собой величину, которая вычисляется на основе двух строк (x, y) одинаковой длины и равна количеству позиций, в которых символы этих строк различны. Таким образом, чем меньше расстояние Хэмминга между признаковыми описаниями двух областей, тем скорее данные области будут обозначать одну и ту же область на изображении.</p> <p>Таким образом вычисляются признаковые описания ключевых точек на изображении с помощью алгоритма BRIEF. Согласно экспериментальным данным, полученным авторами метода, данный алгоритм может достигать точности распознавания в 1,5 раза лучшей по сравнению с аналогом SURF при прочих равных условиях.</p>				
Подпись и дата					
Инв. № дубл.					
Взам. инв. №					
Подпись и дата					
Инв. № подл.					
<p align="center">1.2.2. Методы локализации объектов на изображении</p>					
<p>Данный раздел посвящен методам, использующимся на этапе принятия решения о локализации объектов на изображении на основании признакового описания объекта, полученного на предыдущем этапе. Были подробно рассмотрены следующие алгоритмы:</p>					
Изм.	Лист	№ докум.	Подпись	Дата	<div> <div>ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)</div> <div>Лист</div> <div>16</div> </div>

Перв. примен.	<ul style="list-style-type: none"> • Локализация объекта на основе объединения регионов изображения по принципу схожести [1]. • Иерархическая пирамидальная схема локализации объекта [4]. • Классификатор в виде вероятностных деревьев [6]. <p>Также последний из указанных алгоритмов решает задачу классификации объектов на изображении по категориям.</p>				
Справ. №	<p>Локализация объекта на основе объединения областей изображения по принципу схожести</p> <p>Первый этап метода [1] – выполнение начального разделения исследуемого изображения на области. Это необходимо для произведения дальнейшего объединения полученных областей. Для выполнения данной сегментации используется алгоритм Mean shift.</p> <p>Для каждой полученной области вычисляется цветовая гистограмма, которая является признаком, учитывающим цветовую гамму рассматриваемой области.</p> <p>Важной проблемой является определение меры схожести между областями, которые содержат в себе ключевые точки, т.е. представляют собой искомый объект, и областями, которые ключевых точек не содержат. В качестве такой меры используется известная мера, называемая City Block distance. Исходя из этого мера схожести между двумя областями X и Y вычисляется по формуле 1.8:</p> $\Delta(X, Y) = \sum_{m=1}^{512} Hist_X^m - Hist_Y^m \quad (1.8)$ <p>Низкое значение данной величины означает высокую степень сходства между двумя областями.</p> <p>Правило объединения двух соседних областей звучит следующим образом: пусть X – рассматриваемая область, Y – соседняя с ней область, S_y – набор областей, соседних с Y. В этом случае X и Y должны быть объединены тогда и только тогда, когда степень схожести между ними максимальна по сравнению со всеми степенями схожести между Y и областями из S_y. Данная стратегия объединения областей позволяет избежать объединения областей, относящихся к искомому объекту, с областями, относящимися к фону.</p> <p>В результате итеративного выполнения данного алгоритма выявляется наибольшая область, представляющий собой искомый объект, таким образом решается задача локализации объекта на изображении.</p>				
Подпись и дата					
Инв. № дубл.					
Взам. инв. №					
Подпись и дата					
Инв. № подл.					

Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)	Лист
						17

Перв. примен.					
Справ. №					
Подпись и дата					
Инв. № дубл.					
Взам. инв. №					
Подпись и дата					
Инв. № подл.					
<div> <div>Изм.</div> <div>Лист</div> <div>№ докум.</div> <div>Подпись</div> <div>Дата</div> </div>					<div>Лист</div> <div>18</div>

Иерархическая пирамидальная схема локализации объекта

Данный метод [4] представляет собой сравнительно простой алгоритм локализации объекта на изображении с использованием многоуровневого последовательного деления изображения на равные сегменты. На вход данный алгоритм принимает набор признаковых описаний ключевых точек искомого объекта.

Исследуемое изображение последовательно делится на L уровней, где L – параметр, выбираемый пользователем и контролирующий уровень детализации. На каждом уровне l изображение делится на сегменты с индексами $0 < i < 4^{l-1}$, где $0 < l < L - 1$. На каждом уровне l входной набор признаковых описаний особых точек также делится на сегменты, соответствующие сегментам, полученным на изображении. Индекс каждого сегмента для точки $p = (x, y)$ на изображении вычисляется по формуле 1.9:

$$i = \left[idiv\left(\frac{x-1}{\frac{C}{2^l}}\right) + idiv\left(\frac{y-1}{\frac{R}{2^l}}\right) \right] 2^l + 1 \quad (1.9)$$

где $idiv(x)$ – оператор целочисленного деления, R – максимальное число строк в изображении (высота изображения), C – максимальное число столбцов в изображении (ширина изображения).

На основании такого разбиения изображения на сегменты на каждом уровне пирамиды для каждого сегмента строятся гистограммы на основе значений признаковых описаний, полученных для ключевых точек каждого сегмента. Разбиение изображения на сегменты и пример гистограмм представлен на рисунке 1.6:

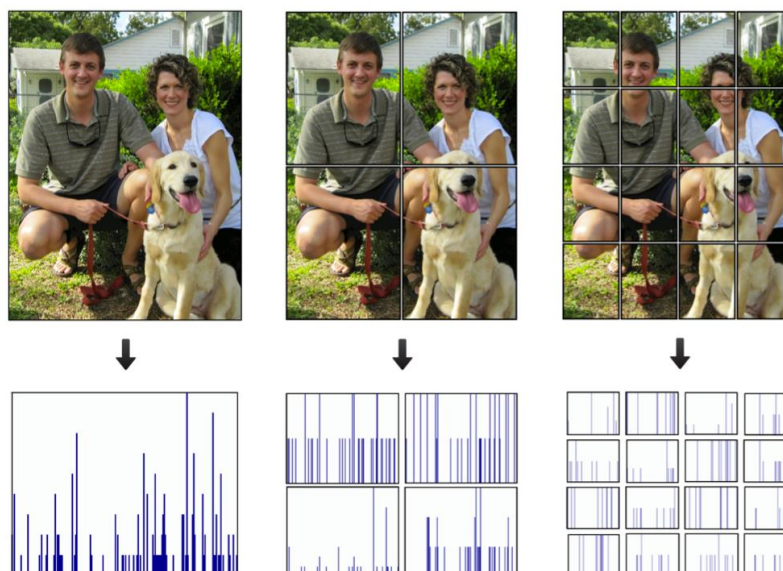


Рисунок 1.6. Разбиение изображения на сегменты и гистограммы полученных сегментов

На финальном этапе данного алгоритма на основе полученных гистограмм делается вывод о том, принадлежит ли каждый из полученных сегментов к искомому объекту. Если гистограмма свидетельствует о наличии ключевых точек в сегменте, то данный сегмент является частью искомого объекта на изображении. В результате этого процесса выявляются и объединяются в единую область все сегменты, относящиеся к объекту, таким образом решается задача локализации объекта на изображении.

Классификатор в виде вероятностных деревьев

Данный классификатор [6] использует дерево, каждый узел которого содержит элементарный тест, делящий пространство классифицируемых данных, в данном случае – пространство областей изображения. Каждый конечный узел – лист – содержит оценку распределения вероятности отнесения области к определенному классу. Данное дерево формируется в процессе обучения алгоритма с использованием обучающей выборки данных.

Каждая новая область изображения, которую необходимо классифицировать, проходит вниз через дерево и на каждом узле подвергается элементарному тесту. По результатам этого теста область отправляется в левый или правый узел на следующий уровень вниз. Данный тест основан на разнице между яркостями двух пикселей, являющихся соседними для ключевой точки:

$$C_2(m_1, m_2) = \begin{cases} \text{левый узел, если } I_\sigma(p, m_1) \leq I_\sigma(p, m_2) \\ \text{правый узел – в остальных случаях} \end{cases} \quad (1.10)$$

где $I_\sigma(p, m)$ – яркость области p в пикселе m .

При достижении листа области присваивается вероятность его принадлежности к определенному классу.

Для достижения наилучшей точности распознавания формируются несколько подобных деревьев (T_1, \dots, T_l) , и алгоритм комбинирует результаты прогона области через все деревья. Увеличение числа деревьев ведет к увеличению точности распознавания за счет использования отличающихся тестов в разных деревьях. В то же время увеличение числа деревьев ведет к увеличению требуемой вычислительной мощности.

Пример дерева и распределения вероятностей в его листьях представлен на рисунке 1.7:

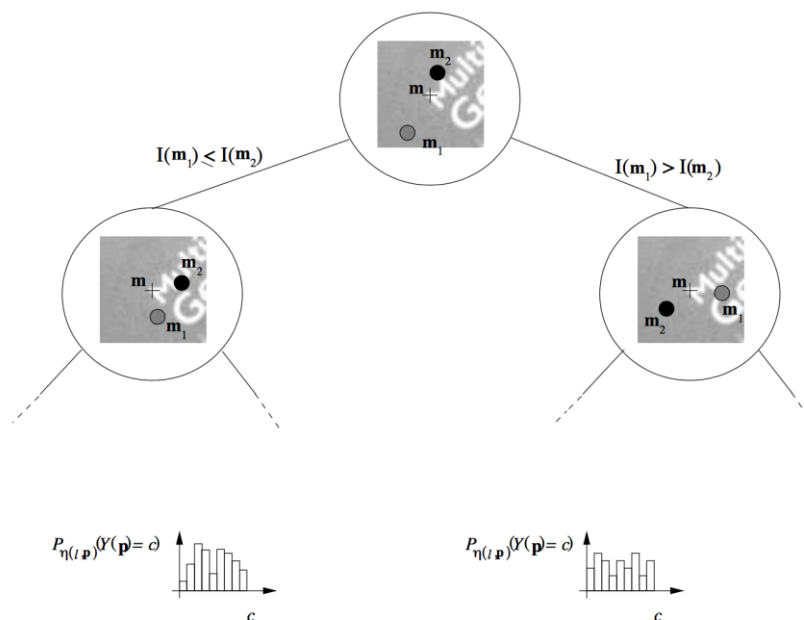


Рисунок 1.7. Дерево элементарных тестов и функции распределения вероятности отнесения области p к определенным классам

Распределение вероятности принадлежности области к определенному классу, содержащееся в листьях деревьев, имеет вид:

$$P_{\eta(l,p)}(Y(p) = c) \quad (1.11)$$

где c – класс объекта, $\eta(l, p)$ – лист дерева T_l , которого достигла область p .

Область p классифицируется на основе средней вероятности:

где $p_c(p)$ – средняя апостериорная вероятность того, что объект, изображенный в области p , относится к классу c . Также вводится предельное значение T_c , определяющее, является ли произведенная классификация корректной или нет:

На практике используется величина s от 60% до 90%. Ключевые точки, для которых $p_c(p) > T_c$, относятся к категории неклассифицируемых и далее исключаются из рассмотрения. Остальные ключевые точки определяются как объекты тех или иных классов. Таким образом решается задача локализации искомого объекта на исследуемом изображении посредством классификации ключевых точек методом вероятностных деревьев.

Перв. примен.	<h3 style="text-align: center;">1.3. Выводы по главе 1</h3> <p>В данной главе представлен обзор некоторых известных методов, используемых при решении задачи локализации объектов на изображении. При составлении данного обзора были выявлены основные подзадачи, которые включает в себя общая задача локализации объектов на изображении. Также были описаны известные способы решения этих подзадач.</p> <p>В ходе написания обзора было обнаружено множество существующих методов, решающих поставленную задачу локализации объекта на изображении, что говорит об актуальности этой задачи в современной теоретической информатике. Также было выявлено, что большинство рассмотренных методов базируются на схожих принципах и отличаются лишь деталями реализации на этапах решения различных подзадач.</p> <p>Одной из целей составления данного обзора было выявление проблем, появляющихся при решении задачи локализации объекта на изображении. При составлении обзора были выявлены следующие проблемы:</p> <ul style="list-style-type: none"> • Проблема формирования признакового описания объекта. Данная проблема связана с выбором системы признаков, которая будет использоваться для однозначного и информативного описания точек искомого объекта. Главная сложность при выборе системы признаков – поиск компромисса между размером признакового описания и его информативностью. Популярные современные алгоритмы [1][2] чаще всего формируют признаковые описания точек объекта в виде векторов чисел с плавающей точкой размерностью 64, 128 и т.д. Такие признаковые описания занимают значительный объем памяти при хранении и требуют высоких вычислительных затрат при своем формировании и последующей обработке. • Проблема поиска ключевых точек. Данная проблема заключается в обнаружении таких точек на объекте, которые однозначно характеризуют данный объект и являются устойчивыми к изменениям масштаба и поворота, а также к шумам и помехам. Для решения этой проблемы необходим подходящий критерий, способный выявить такие ключевые точки на основании их признакового описания. • Проблема принятия решения о локализации объекта. Данная проблема связана с подробным рассмотрением исследуемого изображения. Требуемые вычислительные затраты на данном этапе возрастают по мере увеличения подробности рассмотрения изображения. В то же время на объем необходимых вычислительных затрат оказывает сильное влияние выбранная на 																									
	Справ. №																									
Подпись и дата		Инв. № дубл.	Взам. инв. №	Подпись и дата	Инв. № подл.																					
<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 50%;"></td> <td style="width: 10%; text-align: right;">Лист</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td style="text-align: center;">ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)</td> <td style="text-align: center;">22</td> </tr> <tr> <td>Изм.</td> <td>Лист</td> <td>№ докум.</td> <td>Подпись</td> <td>Дата</td> <td colspan="2"></td> </tr> </table>												Лист						ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)	22	Изм.	Лист	№ докум.	Подпись	Дата		
						Лист																				
					ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)	22																				
Изм.	Лист	№ докум.	Подпись	Дата																						

первом этапе система признаков для формирования признакового описания. Это говорит о значительной важности выбора системы признаков, которая позволит формировать признаковое описание искомого объекта в компактном виде.

На основании выявленных этапов и рассмотренных существующих методов решения задачи локализации объектов на изображении был предложен новый метод решения данной задачи. Данный метод предлагает собственную комбинацию подходов к решению указанных выше проблем. Теоретические основы этого метода решения задачи локализации объекта на изображении будут подробно описаны в главе 2.

Перв. примен.

Справ. №

Подпись и дата

Инв. № дубл.

Взам. инв. №

Подпись и дата

Инв. № подл.

ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)

Лист

23

Изм. Лист № докум. Подпись Дата

Глава 2. Информационная модель локализации объектов на изображении

Предлагаемый метод решения задачи локализации объекта на изображении применяет теорию активного восприятия (ТАВ) [9] на этапе формирования признакового описания изображения и поиска ключевых точек искомого объекта. На этапе принятия решения о локализации объекта на изображении используются методы кластеризации: Mean shift и метод на основе алгоритма кластеризации k-means. Использование нескольких методов кластеризации позволяет добиться более достоверных результатов на этапе принятия решения.

2.1. Теория активного восприятия

Теория активного восприятия (ТАВ) [9] 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.

2.1.1. U-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:

Перв. примен.										
Справ. №										
Подпись и дата										
Инв. № дубл.										
Взам. инв. №										
Подпись и дата										
Инв. № подл.										
					ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)					Лист
					Изм. Лист № докум. Подпись Дата					24

$$D = \{d_i\}, i = \overline{1, M} \quad (2.1)$$

где M – число сегментов, на которое 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 , $k = \overline{1, L}$.

Рассмотрим пример вычисления U-преобразования для одного сегмента изображения с использованием 16 фильтров на этапе дифференцирования.

Первым этапом вычисления U-преобразования является этап интегрирования. 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 2.1:



1209	1446	1027	741
1786	2686	1955	1273
426	688	500	317
288	365	281	243

Рисунок 2.1. Формирование визуальных масс участка изображения

Как видно из рисунка 2.1, более светлым участкам изображения соответствует элемент матрицы визуальных масс с более высоким значением и наоборот.

Далее выполняется этап дифференцирования. В ходе выполнения данного этапа матрица визуальных масс поэлементно умножается на фильтры, представленные на рисунке 2.2. Светлый элемент фильтра соответствует значению «1», темный элемент – значению «-1».

Перв. примен.					
Справ. №					
Подпись и дата					
Инв. № дубл.					
Взам. инв. №					
Подпись и дата					
Инв. № подл.					

F_4

F_{12}

F_5

F_{13}

F_6

F_{14}

F_7

F_{15}

Рисунок 2.2. Фильтры, используемые для вычисления U-преобразования

Например, в результате умножения на фильтр F_8 матрица визуальных масс будет иметь вид, представленный на рисунке 2.3:

					ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)	Лист
						27
Изм.	Лист	№ докум.	Подпись	Дата		

Перв. примен.		Справ. №		Подпись и дата		Инв. № дубл.		Взам. инв. №		Подпись и дата		Инв. № подл.	

1209	-1446	-1027	741
-1786	2686	1955	-1273
-426	688	500	-317
288	-365	-281	243

Рисунок 2.3. Матрица визуальных масс после наложения фильтра F_8

Тогда значением спектрального коэффициента, соответствующего фильтру F_8 , будет являться алгебраическая сумма всех элементов матрицы после наложения данного фильтра. Значение 8-го спектрального коэффициента для данной матрицы равно 1388. Аналогичным образом рассчитываются все 16 спектральных коэффициентов по исследуемому изображению. Результаты вычисления U-преобразования по рассматриваемому изображению приведены в таблице 2.1.

Таблица 2.1. Спектральные коэффициенты U-преобразования

Фильтр	Спектральный коэффициент
0	15232
1	-2556
2	9015
3	-1703
4	2667
5	4030
6	1547
7	-526
8	1388
9	-285
10	2524
11	-50
12	-191
13	-308
14	730
15	-228

Изм.	Лист	№ докум.	Подпись	Дата

ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)

Лист
28

В результате применения U-преобразования к сегменту изображения были получены 16 спектральных коэффициентов (по числу используемых фильтров). Данные спектральные коэффициенты могут быть использованы как признаковое описание данного сегмента.

2.1.2. 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 2.2. 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 2.4:

-1	-1	1	1
-1	-1	1	1
1	1	-1	-1
1	1	-1	-1

→

0	0	1	1
0	0	1	1
1	1	0	0
1	1	0	0

Рисунок 2.4. 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

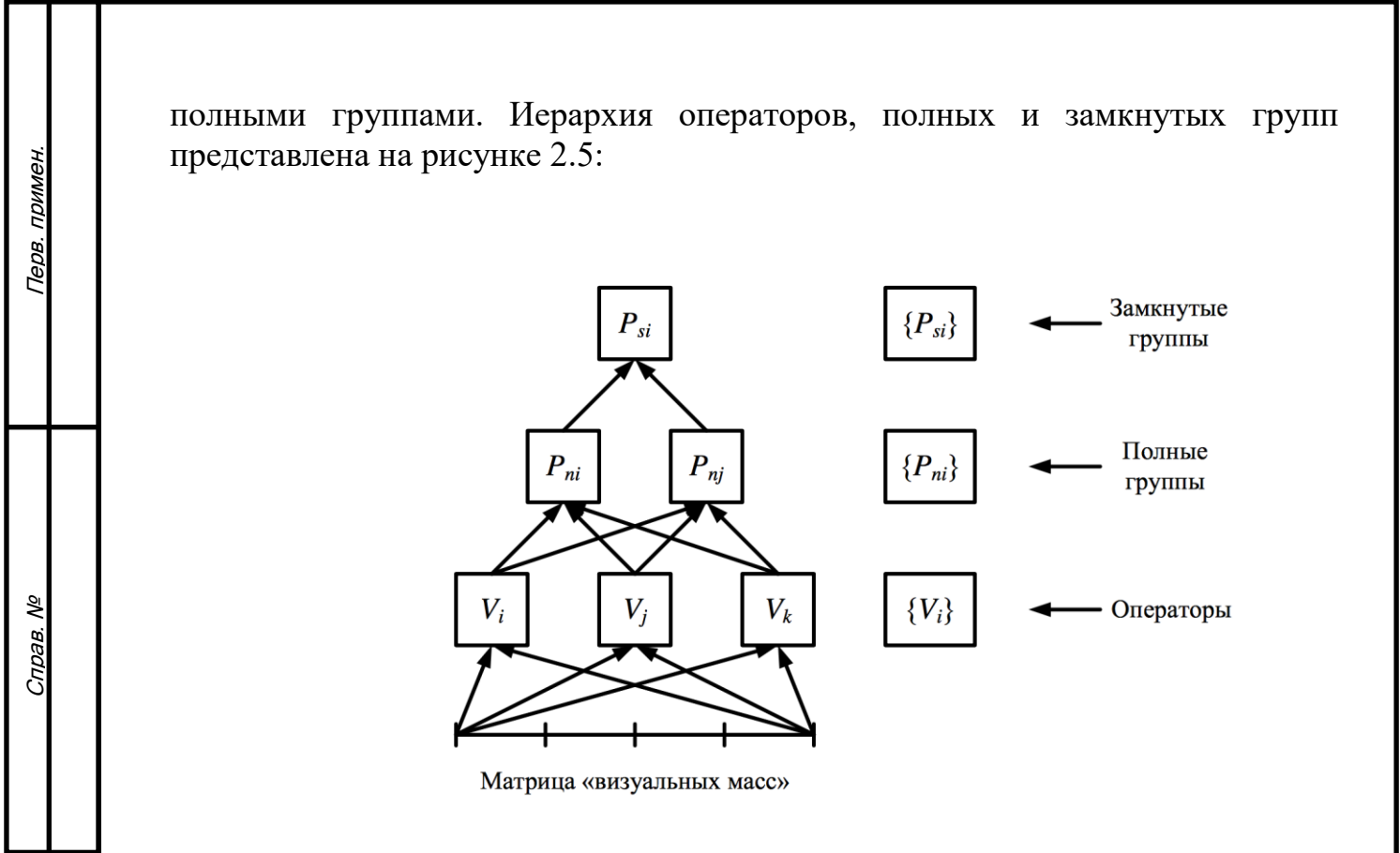


Рисунок 2.5. Иерархическая структура операторов, полных и замкнутых групп

Подпись и дата

Инв. № дубл.

Взам. инв. №

Подпись и дата

Инв. № подл.

Каждому оператору или группе ставится в соответствие понятие образа оператора или группы. Образ оператора V_3 представлен на рисунке 2.4. На основании подобных образов происходит вычисление масс операторов, полных и замкнутых групп. Пример вычисления массы замкнутой группы по сегменту изображения (рисунок 2.1) представлен на рисунке 2.6:

Σ

1209	1446	1027	741
1786	2686	1955	1273
426	688	500	317
288	365	281	243

 \times

0	0	1	1
0	0	1	1
0	0	1	1
1	1	0	0

 $= 6466$

Рисунок 2.6. Вычисление массы замкнутой группы по сегменту изображения

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)

Лист
30

Перв. примен.					
Справ. №					
<p>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 точек исходного изображения.</p>					
Подпись и дата					
Инв. № дубл.					
Взам. инв. №					
Подпись и дата					
Инв. № подл.					
Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)
					Лист
					31

2.2. Реализация системы локализации объекта на изображении на основе теории активного восприятия

Теория активного восприятия предоставляет обширный инструментарий для решения задач описания и последующего анализа изображений. Данный раздел посвящен деталям реализации системы локализации объекта на изображении с помощью теории активного восприятия на этапе формирования признакового описания изображения и поиска ключевых точек объекта. Также в данном разделе описываются варианты реализации этапа принятия решения о локализации объекта на изображении.

Информационная модель решения задачи локализации объекта на изображении с помощью теории активного восприятия представлена на рисунке 2.7:

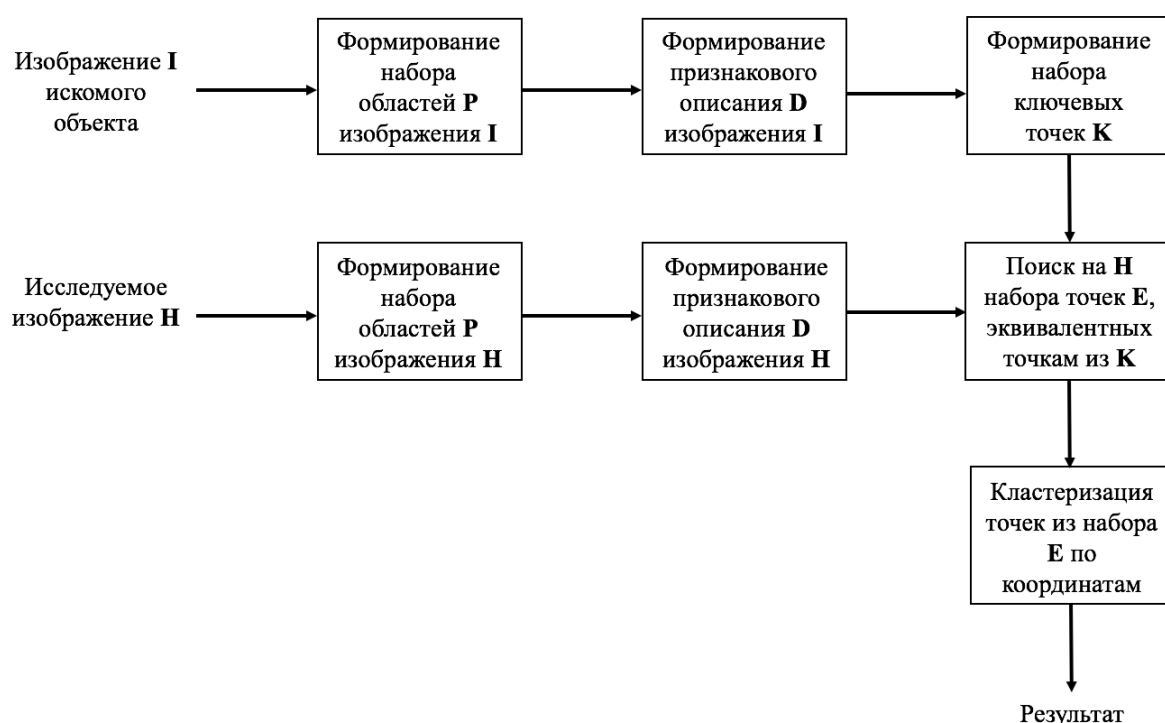


Рисунок 2.7. Информационная модель решения задачи локализации объектов на изображении

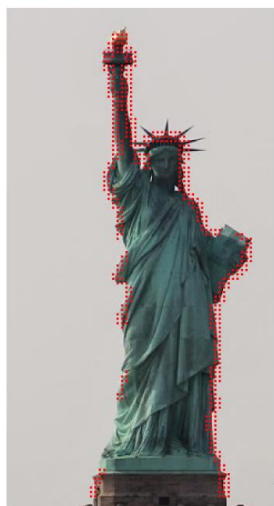
Далее представлено подробное описание каждого этапа, обозначенного на данной информационной модели.

2.2.1. Предварительная обработка изображения

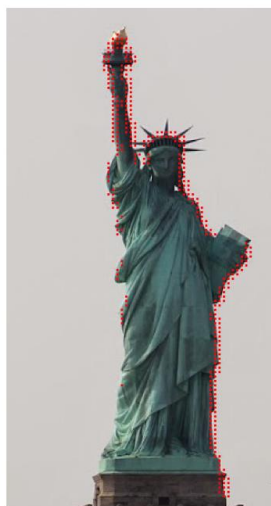
Пусть изображение искомого объекта закодировано согласно модели RGB. В этом случае каждый пиксель изображения представляется в виде

Перв. примен.	<p>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:</p> $I(x, y) = \frac{(R_{xy} + G_{xy} + B_{xy})}{3}, x \in X, y \in Y \quad (2.2)$ <p>где R_{xy}, G_{xy}, B_{xy} – яркости соответствующих компонент модели RGB для пикселя с координатами (x, y); X, Y – множества соответствующих координат изображения.</p> <p>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. Путем прохождения по данной матрице скользящим окном размера $n \times n$ элементов можно представить изображение в виде набора областей:</p> $P = \{P_{ij}\}, i \in X, j \in Y \quad (2.3)$ <p>где P_{ij} – область изображения размером $n \times n$ пикселей; i, j – координаты левого верхнего угла области P_{ij}; n – размер области P_{ij}.</p> <p>Для достижения большей степени детализации признакового описания изображения набор областей формируется с «перекрытием», т.е. разница между координатами (i, j) двух соседних областей по X и по Y (смещение) меньше, чем размер области n.</p> <p>Таким образом, изображение искомого объекта представляется в виде набора областей размера $n \times n$. Данное представление является подходящим для выполнения следующего этапа решения задачи – формирования признакового описания.</p> <p>2.2.2. Формирование признакового описания искомого объекта</p> <p>Признаковое описание искомого объекта представляет собой набор признаковых описаний областей данного объекта, полученных на предыдущем этапе. В качестве описания областей объекта на данном этапе были выбраны коэффициенты спектрального представления изображения, получаемые в результате выполнения U-преобразования.</p> <p>Для формирования признакового описания для каждой полученной области изображения вычисляется U-преобразование. Коэффициенты спектрального представления, полученные в результате выполнения U-преобразования, и являются признаковым описанием данной области</p>																	
	Справ. №																	
Подпись и дата		Инв. № дубл.	Взам. инв. №	Подпись и дата	Инв. № подл.													
<table border="1"> <tr> <td>Изм.</td> <td>Лист</td> <td>№ докум.</td> <td>Подпись</td> <td>Дата</td> <td rowspan="2">ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)</td> <td>Лист</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td>33</td> </tr> </table>						Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)	Лист						33
Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)	Лист												
						33												

Перв. примен.	<p>изображения. Таким образом, признаковое описание области изображения P_{ij} представляется в виде:</p> $D_{ij} = \{d_k\}, k = \overline{1, L} \quad (2.4)$ <p>где d_k – k-й спектральный коэффициент спектрального представления области изображения; L – количество фильтров, применяемых при вычислении U-преобразования.</p> <p>Исходя из этого признаковое описание изображения в целом можно представить в виде набора признаковых описаний всех его областей:</p> $D = \{D_{ij}\}, i \in X, j \in Y \quad (2.5)$ <p>Данное признаковое описание является достаточно ёмким, чтобы однозначно описать каждую область изображения искомого объекта. При этом его размер (16 чисел с плавающей точкой на одну точку изображения) не требует высоких затрат памяти при хранении.</p> <p>Признаковое описание искомого объекта может быть существенно сокращено путем выделения особых точек, характерных для данного объекта. Сокращение признакового описания происходит за счет удаления из него областей, которые не содержат ключевых точек.</p> <p>В данной работе точка изображения считается ключевой, если в области, в которой находится эта точка, имеется резкий перепад яркости. Перепады яркости соответствуют контурам искомого объекта. Признаковое описание областей изображения, состоящее из спектральных коэффициентов U-преобразования, позволяет легко выявлять области с перепадами яркости внутри.</p> <p>Как следует из формулы 2.4, каждая область изображения представлена в виде набора спектральных коэффициентов. Для каждой области вычисляется среднеквадратичное отклонение спектральных коэффициентов. При этом коэффициент, соответствующий отклику на фильтр F_0 (сумма всех визуальных масс области), не учитывается при вычислении СКО. Исходя из этого выводится формула 2.6 для вычисления СКО спектральных коэффициентов области P_{ij}:</p>																							
	Справ. №																							
Подпись и дата		Инв. № дубл.	Взам. инв. №	Подпись и дата	Инв. № подл.																			
<table border="1"> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td rowspan="3">ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)</td> <td>Лист</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td>34</td> </tr> <tr> <td>Изм.</td> <td>Лист</td> <td>№ докум.</td> <td>Подпись</td> <td>Дата</td> <td></td> </tr> </table>											ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)	Лист						34	Изм.	Лист	№ докум.	Подпись	Дата	
					ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)	Лист																		
						34																		
Изм.	Лист	№ докум.	Подпись	Дата																				



$k = 0,6$
 $N = 517$



$k = 0,7$
 $N = 351$



$k = 0,8$
 $N = 173$

Рисунок 2.8. Примеры выявления областей изображения с ключевыми точками при различных значениях коэффициента k

Как видно из рисунка 2.8, увеличение коэффициента отбора k приводит к уменьшению количества ключевых точек, выявленных на рассматриваемом объекте. Изменение данного коэффициента позволяет между тщательностью выделение контуров объекта (количеством ключевых точек) и объемом памяти, необходимым для их хранения. Важно отметить, что при любом значении k в категорию ключевых точек попадают только те точки изображения, области вокруг которых содержат в себе перепад яркости, т.е. контур.

Таким образом, выявление ключевых точек объекта позволяет существенно сократить число областей в признаковом описании. Это приводит к значительному снижению затрат памяти и вычислительной мощности на этапе принятия решения о локализации объекта.

2.2.3. Принятие решения о локализации объекта на изображении

Данный этап заключается в рассмотрении исследуемого изображения и выявлении на нем точек, эквивалентных по своему признаковому описанию ключевым точкам искомого объекта, полученным на предыдущем этапе. На основе этого происходит принятие решения о наличии или отсутствии искомого объекта на рассматриваемом изображении, и в случае его наличия выполняется локализация.

Для выявления точек, эквивалентных ключевым точкам по признаковому описанию, вводится понятие критерия эквивалентности точек

Перв. примен.	<p>изображения. Пусть две точки P_1 и P_2 имеют признаковые описания D_1 и D_2. Тогда критерий эквивалентности точек P_1 и P_2 имеет следующий вид:</p> $\text{sign}(D_{1i}) = \text{sign}(D_{2i}), i = \overline{1, L} \quad (2.9)$ <p>где $\text{sign}(x)$ – оператор вычисления знака выражения x; D_{1i}, D_{2i} – i-й элемент вектора спектральных коэффициентов точек P_1 и P_2; L – количество спектральных коэффициентов в признаковых описаниях точек ($L = 16$).</p> <p>Таким образом, две точки изображения считаются эквивалентными тогда и только тогда, когда соответствующие спектральные коэффициенты признаковых описаний данных точек имеют одинаковые знаки.</p> <p>После обнаружения на изображении точек, эквивалентных ключевым точкам объекта, необходимо принять решение о локализации данного объекта в месте наибольшего скопления найденных точек. В данной работе предлагается метод принятия такого решения на основе разбиения точек на кластеры по координатному принципу. В таком случае делается предположение, что наибольший по количеству точек кластер располагается на изображении в то же месте, где и искомый объект.</p> <p>В настоящее время существует множество методов, применяемых для решения задачи кластеризации объектов [10]. Все они могут давать различные результаты на одних и тех же входных данных, однако сравнить качество работы методов крайне сложно, так как не существует однозначно наилучшего критерия оценки качества кластеризации.</p> <p>Для получения более высокой точности результатов в данной работе используются два различных метода кластеризации точек:</p> <ul style="list-style-type: none"> • Метод на основе алгоритма кластеризации k-means [11]. • Метод кластеризации Mean shift [12][13]. <p>Рассмотрим подробнее данные методы.</p> <p>Метод кластеризации на основе алгоритма k-means</p> <p>Алгоритм кластеризации k-means (k-средних) [11] является одним из наиболее популярных алгоритмов кластеризации. Он был предложен в 1950-х годах одновременно Г. Штейнгуазом и С. Ллойдом.</p> <p>Данный алгоритм разбивает множество объектов на заранее известное число кластеров k.</p> <p>На первом шаге алгоритма для каждого кластера случайным образом выбираются центры масс. После этого все объекты разбиваются на кластеры по принципу близости к центрам масс кластеров. На следующих итерациях вычисляются новые центры масс на основе кластеров, полученных на предыдущем этапе.</p>																	
	Справ. №																	
Подпись и дата		Инв. № дубл.	Взам. инв. №	Подпись и дата	Инв. № подл.													
<table border="1"> <tr> <td>Изм.</td> <td>Лист</td> <td>№ докум.</td> <td>Подпись</td> <td>Дата</td> <td rowspan="2">ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)</td> <td>Лист</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td>37</td> </tr> </table>						Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)	Лист						37
Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)	Лист												
						37												

Перв. примен.		<p>В процессе своей работы алгоритм стремится минимизировать суммарное квадратичное отклонение точек кластеров от центров этих кластеров:</p> $V = \sum_{i=1}^k \sum_{x_j \in S_i} (x_j - \mu_i)^2 \quad (2.10)$ <p>где k – число кластеров; S_i – i-й кластер, $i = \overline{1, k}$; μ_i – центр масс кластера S_i; x_j – j-й элемент кластера S_i.</p> <p>Алгоритм завершает свою работу, когда на одной из итераций не происходит изменение центров масс кластеров. Работа данного алгоритма всегда завершается за конечное число шагов.</p> <p>Однако алгоритм k-means имеет ряд недостатков:</p> <ul style="list-style-type: none"> • Не гарантируется достижение глобального минимума величины V. • Результат зависит от выбора начальных центров масс кластеров, оптимальный способ их выбора неизвестен. • Число кластеров задается как входной параметр алгоритма. <p>Именно последний из указанных недостатков приводит к тому, что использование стандартной версии алгоритма k-means не позволяет решить задачу локализации точек объекта на изображении, так как заранее неизвестно число кластеров, на которые необходимо разделить точки. Поэтому в данной работе была предложена доработанная версия метода кластеризации на основе алгоритма k-means.</p> <p>Предложенный метод кластеризации итеративно выполняет кластеризацию с помощью алгоритма k-means с входными параметрами $k = \overline{1, N}$, где N – количество точек, которые необходимо разделить на кластеры. На каждой итерации метода выявляется кластер, наибольший по количеству входящих в него точек. Для этого кластера вычисляется сумма расстояний между всеми точками, входящими в кластер:</p> $s = \sum_{i=1}^M \sum_{j=1}^M dist(p_i, p_j), i \neq j \quad (2.11)$ <p>где M – размер рассматриваемого кластера; p_i, p_j – точки кластера, между которыми вычисляется расстояние; $dist(x, y)$ – функция вычисления Евклидова расстояния между двумя точками по их координатам.</p> <p>С учетом суммы расстояний между точками внутри кластера записывается критерий остановки работы метода:</p>				
Справ. №						
Подпись и дата						
Инв. № дубл.						
Взам. инв. №						
Подпись и дата						
Инв. № подл.						
Изм.	Лист	№ докум.	Подпись	Дата	<div>ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)</div> <div>Лист 38</div>	

$$\frac{s}{(M-1)^2} < eps \quad (2.12)$$

где s – сумма расстояний между точками внутри кластера; M – размер кластера; eps – входной параметр метода, в данной работе этот параметр определяется как треть диагонали изображения искомого объекта.

Таким образом, метод кластеризации на основе алгоритма k-means позволяет разделять точки изображения на заранее неизвестное количество кластеров в зависимости от заданных входных параметров. Пример результата работы данного метода представлен на рисунке 2.9:



Рисунок 2.9. Результат кластеризации обнаруженных особых точек с помощью метода на основе алгоритма k-means

Метод кластеризации Mean shift

Метод кластеризации Mean shift [12] был предложен К. Фукунагой и Л. Хостетлером в 1975 году. Данный метод является непараметрическим методом кластеризации, который не использует заранее подготовленной информации о количестве кластеров и не ограничивает форму кластеров.

Для каждого набора точек вычисляется функция плотности ядра:

$$f(x) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right) \quad (2.13)$$

где n – количество точек; d – количество измерений пространства данных точек; h – радиус ядра; $K(x)$ – ядро, которое вычисляется по формуле 2.14:

$$K(x) = c_{k,d} k(\|x\|^2) \quad (2.14)$$

где $c_{k,d}$ – константа нормализации.

Исходя из этого вычисляется вектор mean shift, который всегда направлен в направлении максимального увеличения функции плотности $f(x)$.

Алгоритм кластеризации Mean shift является практическим применением процесса поиска максимума функции плотности. Набор всех точек в пространстве, которые сходятся к одному и тому же максимуму функции плотности, называется множеством точек данного максимума. Считается, что точки, попавшие в одно и то же множество, относятся к одному и тому же кластеру.

Таким образом, алгоритм кластеризации Mean shift позволяет разделять точки изображения на заранее неизвестное число кластеров. Пример результата работы данного метода представлен на рисунке 2.10:

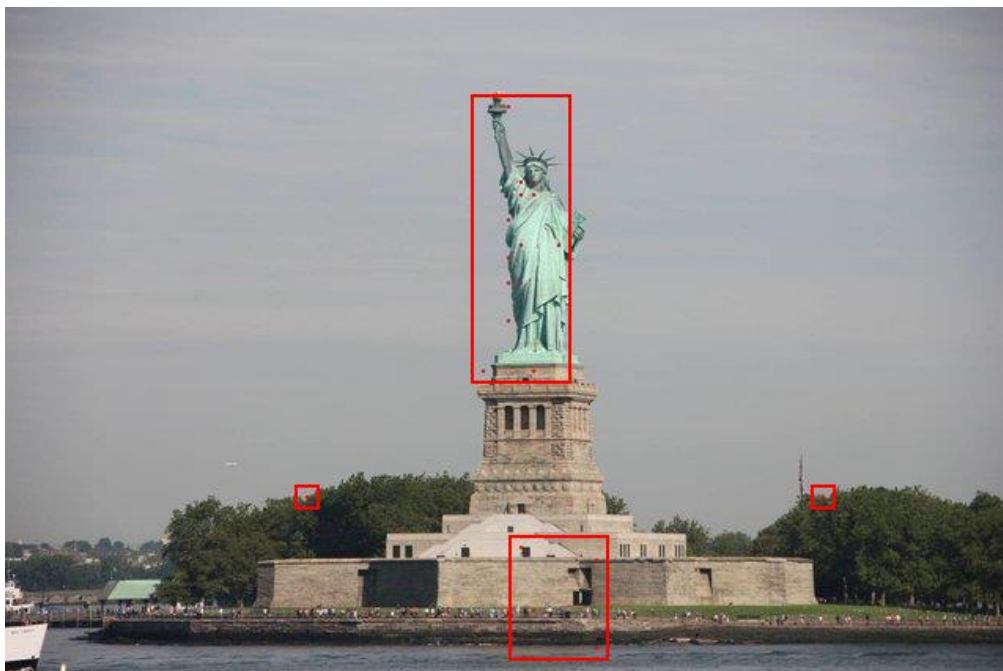


Рисунок 2.10. Результат кластеризации обнаруженных особых точек с помощью метода Mean shift

Изм.	Лист	№ докум.	Подпись	Дата

Перв. примен.						
Справ. №						
Подпись и дата						
Инв. № дубл.						
Взам. инв. №						
Подпись и дата						
Инв. № подл.						
Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)	Лист
						41

Как видно из рисунков 2.9 и 2.10, два рассмотренных метода кластеризации дают различные результаты при одинаковых исходных данных. Однако, кластер, содержащий наибольшее количество точек, соответствующий расположению искомого объекта на изображении, выявлен одинаково обоими методами. Это подтверждает правильность работы каждого из методов.

Перв. примен.																			
Справ. №																			
<h3 style="text-align: center;">2.3. Выводы по главе 2</h3> <p>В данной главе были рассмотрены теоретические основы метода локализации объекта на изображении на основе теории активного восприятия. Были предложены варианты реализации каждого из этапов решения задачи локализации объекта на изображении:</p> <ul style="list-style-type: none"> • Предварительная обработка изображения – разбиение изображения на области. • Формирование признакового описания искомого объекта – теория активного восприятия (U-преобразование). • Принятие решения о локализации объекта – методы кластеризации (k-means, Mean shift). <p>Использование данного набора вариантов реализации каждого из этапов решения задачи локализации 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.</p>																			
Инв. № подл.																			
Подпись и дата																			
Взам. инв. №																			
Инв. № дубл.																			
Подпись и дата																			
<table border="1" style="width: 100%;"> <tr> <td>Изм.</td> <td>Лист</td> <td>№ докум.</td> <td>Подпись</td> <td>Дата</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> </table>					Изм.	Лист	№ докум.	Подпись	Дата						<table border="1" style="width: 100%;"> <tr> <td colspan="2">ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)</td> </tr> <tr> <td>Лист</td> <td>42</td> </tr> </table>	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)		Лист	42
Изм.	Лист	№ докум.	Подпись	Дата															
ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)																			
Лист	42																		

Глава 3. Вычислительный эксперимент

Вычислительный эксперимент является неотъемлемой частью исследования по разработке метода решения задачи локализации объектов на изображении. Результаты эксперимента позволяют сделать вывод о том, насколько успешно предложенный метод справляется с решением задачи по сравнению с известными аналогами, необходимы ли методу доработки, имеет ли смысл применять его 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. В базе содержатся изображения 100 различных объектов, в качестве которых взяты известные архитектурные сооружения и памятники. Каждый объект в базе имеет четыре шаблона, которые и используются при локализации этого объекта на изображении. Первый из шаблонов объекта представляет собой его изображение без добавления шумов, тестирование алгоритма на таких шаблонах далее будет называться тестированием в нормальных условиях. Остальные шаблоны объекта являются его изображениями, полученными в результате наложения на первый шаблон шумов различной интенсивности. Примеры шаблонов одного объекта с различным уровнем шумов представлены на рисунке 3.1:



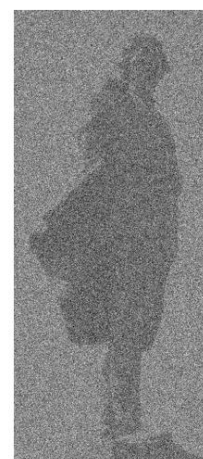
Без добавления
шума



ОСШ = 20 дБ



ОСШ = 10 дБ



ОСШ = 0 дБ

Рисунок 3.1. Шаблоны одного объекта с различным уровнем шумов

Перв. примен.					
Справ. №					
Подпись и дата					
Инв. № дубл.					
Взам. инв. №					
Подпись и дата					
Инв. № подл.					

Собранная база содержит по 100 изображений каждого объекта. Часть этих изображений имеют такой же масштаб и угол поворота объекта, как и шаблоны. Другая часть изображает объект с отличными от шаблонов масштабом или углом поворота. Такие изображения используются для проверки устойчивости разработанного метода к изменениям масштаба и поворота объекта. Таким образом, ёмкость собранной базы составляет 400 шаблонов объектов и 10000 изображений.

Все изображения в базе имеют формат PNG и размер, не превышающий 1000 x 1000 пикселей. Также для каждого изображения имеется файл-описатель формата XML. Данный файл содержит координаты углов прямоугольника, в котором заключен искомый объект на данном изображении.

Предлагаемый метод локализации объекта на изображении в качестве выходных данных выдает координаты углов прямоугольника, в котором предположительно находится искомый объект на исследуемом изображении. Результатом сравнения координат прямоугольника, взятого из файла-описателя изображения, и прямоугольника, полученного в результате работы алгоритма, является вывод о том, верно ли был локализован объект на изображении. Возможны следующие варианты результата обработки предлагаемым алгоритмом одного изображения:

1. Объект локализован на изображении в верной позиции.
2. Объект не локализован на изображении, на котором он отсутствует.
3. Объект локализован на изображении в позиции, в которой он отсутствует.
4. Объект не локализован на изображении, на котором он присутствует.

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.

Предложенный метод решения задачи локализации объектов на изображении имеет определенный набор входных параметров. Различные значения параметров в этом наборе могут существенно влиять на результат работы алгоритма. В ходе эксперимента были исследованы различные наборы значений входных параметров алгоритма с целью выявления набора, позволяющего производить локализацию объектов с наибольшей точностью. Были исследованы следующие входные параметры алгоритма и их возможные значения:

- Коэффициент отбора ключевых точек, k. Значения: 0,6; 0,65; 0,7.
- Размер квадратных областей, на которые разбивается изображение, n. Значения: 16; 32; 64.
- Смещение разбиения изображения на области, m. Значения: 4; 8.

Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)	Лист
						44

Перв. примен.	<h3>3.2. Описание программного продукта</h3> <p>Для проведения описанного выше вычислительного эксперимента был разработан программный продукт на языке программирования R. Данный язык был выбран, так как он обладает широким инструментарием для обработки статистических данных и работы с графикой, а также высокую степень расширяемости с помощью пакетов.</p> <p>Далее рассмотрены основные модули программы, реализующие локализацию объекта на изображении:</p> <ol style="list-style-type: none"> 1. <code>getKeyPoints ()</code> – функция для поиска ключевых точек на изображении. <p style="text-align: center;">Таблица 3.1. Описание функции <code>getKeyPoints ()</code>.</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <th>№</th> <th>Имя параметра</th> <th>Тип параметра</th> <th>Комментарий</th> </tr> <tr> <td>1</td> <td><code>picture</code></td> <td>Входной</td> <td>Исследуемое изображение</td> </tr> <tr> <td>2</td> <td><code>outputPath</code></td> <td>Входной</td> <td>Путь к файлу изображения с найденными ключевыми точками</td> </tr> <tr> <td>3</td> <td><code>coef</code></td> <td>Входной</td> <td>Коэффициент отбора точек, k</td> </tr> <tr> <td>4</td> <td><code>windowSize</code></td> <td>Входной</td> <td>Размер области, n</td> </tr> <tr> <td>5</td> <td><code>step</code></td> <td>Входной</td> <td>Шаг разбиения на области, m</td> </tr> <tr> <td>6</td> <td><code>points</code></td> <td>Выходной</td> <td>Список ключевых точек</td> </tr> </table> <ol style="list-style-type: none"> 2. <code>getKMeansClusters ()</code> – функция для разбиения точек на кластеры с помощью алгоритма k-means. <p style="text-align: center;">Таблица 3.2. Описание функции <code>getKMeansClusters ()</code>.</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <th>№</th> <th>Имя параметра</th> <th>Тип параметра</th> <th>Комментарий</th> </tr> <tr> <td>1</td> <td><code>points</code></td> <td>Входной</td> <td>Точки, подлежащие разделению на кластеры</td> </tr> <tr> <td>2</td> <td><code>eps</code></td> <td>Входной</td> <td>Величина, необходимая для остановки алгоритма</td> </tr> <tr> <td>3</td> <td><code>clusters</code></td> <td>Выходной</td> <td>Список полученных кластеров</td> </tr> </table>					№	Имя параметра	Тип параметра	Комментарий	1	<code>picture</code>	Входной	Исследуемое изображение	2	<code>outputPath</code>	Входной	Путь к файлу изображения с найденными ключевыми точками	3	<code>coef</code>	Входной	Коэффициент отбора точек, k	4	<code>windowSize</code>	Входной	Размер области, n	5	<code>step</code>	Входной	Шаг разбиения на области, m	6	<code>points</code>	Выходной	Список ключевых точек	№	Имя параметра	Тип параметра	Комментарий	1	<code>points</code>	Входной	Точки, подлежащие разделению на кластеры	2	<code>eps</code>	Входной	Величина, необходимая для остановки алгоритма	3	<code>clusters</code>	Выходной	Список полученных кластеров
	№	Имя параметра	Тип параметра	Комментарий																																													
1	<code>picture</code>	Входной	Исследуемое изображение																																														
2	<code>outputPath</code>	Входной	Путь к файлу изображения с найденными ключевыми точками																																														
3	<code>coef</code>	Входной	Коэффициент отбора точек, k																																														
4	<code>windowSize</code>	Входной	Размер области, n																																														
5	<code>step</code>	Входной	Шаг разбиения на области, m																																														
6	<code>points</code>	Выходной	Список ключевых точек																																														
№	Имя параметра	Тип параметра	Комментарий																																														
1	<code>points</code>	Входной	Точки, подлежащие разделению на кластеры																																														
2	<code>eps</code>	Входной	Величина, необходимая для остановки алгоритма																																														
3	<code>clusters</code>	Выходной	Список полученных кластеров																																														
Справ. №																																																	

Подпись и дата					
Инв. № дубл.					
Взам. инв. №					
Подпись и дата					
Инв. № подл.					

					ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)	Лист
Изм.	Лист	№ докум.	Подпись	Дата		46

3. `getMeanShiftClusters ()` – функция для разбиения точек на кластеры с помощью алгоритма Mean shift.

Таблица 3.3. Описание функции `getMeanShiftClusters ()`.

№	Имя параметра	Тип параметра	Комментарий
1	points	Входной	Точки, подлежащие разделению на кластеры
2	eps	Входной	Величина, необходимая для останова алгоритма
3	clusters	Выходной	Список полученных кластеров

4. `isEqual ()` – функция для выявления эквивалентности между областями изображения.

Таблица 3.4. Описание функции `isEqual ()`.

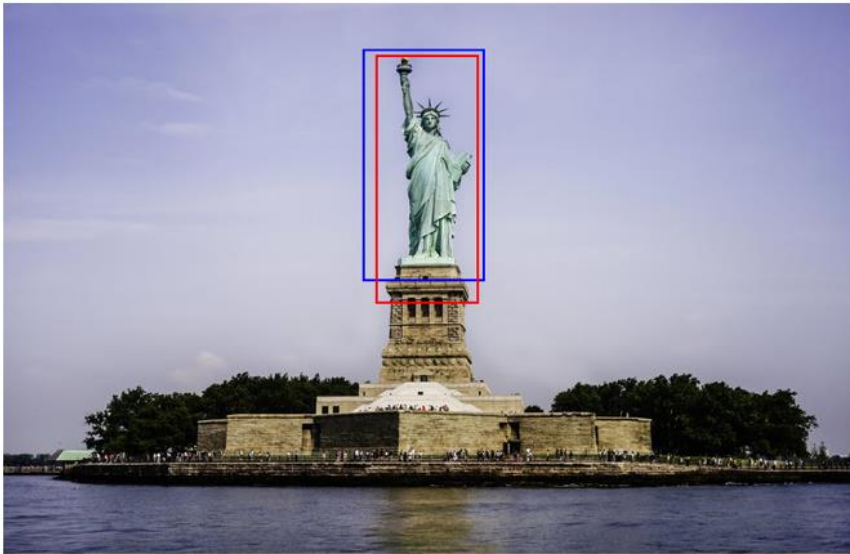

№	Имя параметра	Тип параметра	Комментарий
1	region1	Входной	Первая область
2	region2	Входной	Вторая область
3	result	Выходной	Эквивалентность областей (true/false)

5. `getIntersection ()` – функция для вычисления площади пересечения прямоугольника, содержащего искомое изображение, и прямоугольника, содержащего обнаруженный кластер.

Таблица 3.5. Описание функции `getIntersection ()`.

№	Имя параметра	Тип параметра	Комментарий
1	rect	Входной	Прямоугольник, содержащий искомое изображение
2	cluster	Входной	Кластер, полученный в ходе работы алгоритма

Перв. примен.	<p style="text-align: center;">Таблица 3.5. Описание функции <code>getIntersection ()</code>.</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <th style="width: 5%;">№</th> <th style="width: 25%;">Имя параметра</th> <th style="width: 20%;">Тип параметра</th> <th style="width: 50%;">Комментарий</th> </tr> <tr> <td style="text-align: center;">3</td> <td><code>intersection</code></td> <td>Выходной</td> <td>Площадь пересечения прямоугольника и кластера</td> </tr> </table>					№	Имя параметра	Тип параметра	Комментарий	3	<code>intersection</code>	Выходной	Площадь пересечения прямоугольника и кластера																																				
	№	Имя параметра	Тип параметра	Комментарий																																													
3	<code>intersection</code>	Выходной	Площадь пересечения прямоугольника и кластера																																														
Справ. №	<p>6. <code>findImage ()</code> – функция для поиска объекта на изображении.</p> <p style="text-align: center;">Таблица 3.6 Описание функции <code>findImage ()</code>.</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <th style="width: 5%;">№</th> <th style="width: 25%;">Имя параметра</th> <th style="width: 20%;">Тип параметра</th> <th style="width: 50%;">Комментарий</th> </tr> <tr> <td style="text-align: center;">1</td> <td><code>picture</code></td> <td>Входной</td> <td>Исследуемое изображение</td> </tr> <tr> <td style="text-align: center;">2</td> <td><code>query</code></td> <td>Входной</td> <td>Изображение искомого объекта</td> </tr> <tr> <td style="text-align: center;">3</td> <td><code>keyPoints</code></td> <td>Входной</td> <td>Список ключевых точек искомого объекта</td> </tr> <tr> <td style="text-align: center;">4</td> <td><code>rect</code></td> <td>Входной</td> <td>Прямоугольник, содержащий искомый объект на изображении</td> </tr> <tr> <td style="text-align: center;">5</td> <td><code>outputPath</code></td> <td>Входной</td> <td>Путь к файлу изображения с локализованным объектом</td> </tr> <tr> <td style="text-align: center;">6</td> <td><code>coef</code></td> <td>Входной</td> <td>Коэффициент отбора точек, k</td> </tr> <tr> <td style="text-align: center;">7</td> <td><code>windowSize</code></td> <td>Входной</td> <td>Размер области, n</td> </tr> <tr> <td style="text-align: center;">8</td> <td><code>step</code></td> <td>Входной</td> <td>Шаг разбиения на области, m</td> </tr> <tr> <td style="text-align: center;">9</td> <td><code>clusterization</code></td> <td>Входной</td> <td>Используемый метод кластеризации</td> </tr> <tr> <td style="text-align: center;">10</td> <td><code>result</code></td> <td>Выходной</td> <td>Результат локализации изображения (успех/ошибка 1-го рода/ошибка 2-го рода)</td> </tr> </table>					№	Имя параметра	Тип параметра	Комментарий	1	<code>picture</code>	Входной	Исследуемое изображение	2	<code>query</code>	Входной	Изображение искомого объекта	3	<code>keyPoints</code>	Входной	Список ключевых точек искомого объекта	4	<code>rect</code>	Входной	Прямоугольник, содержащий искомый объект на изображении	5	<code>outputPath</code>	Входной	Путь к файлу изображения с локализованным объектом	6	<code>coef</code>	Входной	Коэффициент отбора точек, k	7	<code>windowSize</code>	Входной	Размер области, n	8	<code>step</code>	Входной	Шаг разбиения на области, m	9	<code>clusterization</code>	Входной	Используемый метод кластеризации	10	<code>result</code>	Выходной	Результат локализации изображения (успех/ошибка 1-го рода/ошибка 2-го рода)
	№	Имя параметра	Тип параметра	Комментарий																																													
1	<code>picture</code>	Входной	Исследуемое изображение																																														
2	<code>query</code>	Входной	Изображение искомого объекта																																														
3	<code>keyPoints</code>	Входной	Список ключевых точек искомого объекта																																														
4	<code>rect</code>	Входной	Прямоугольник, содержащий искомый объект на изображении																																														
5	<code>outputPath</code>	Входной	Путь к файлу изображения с локализованным объектом																																														
6	<code>coef</code>	Входной	Коэффициент отбора точек, k																																														
7	<code>windowSize</code>	Входной	Размер области, n																																														
8	<code>step</code>	Входной	Шаг разбиения на области, m																																														
9	<code>clusterization</code>	Входной	Используемый метод кластеризации																																														
10	<code>result</code>	Выходной	Результат локализации изображения (успех/ошибка 1-го рода/ошибка 2-го рода)																																														
Подпись и дата	<p>В результате обработки каждого изображения из базы данным программным продуктом формируется выходное изображение. Выходное изображение содержит на себе обнаруженные точки, эквивалентные ключевым, позицию искомого объекта, и найденный кластер, соответствующий локализованному объекту. Примеры результата работы</p>																																																
	Инов. № подл.	<div style="display: flex; justify-content: space-between; align-items: flex-end;"> <div style="width: 40%;"> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 15%;">Изм.</td> <td style="width: 15%;">Лист</td> <td style="width: 15%;">№ докум.</td> <td style="width: 15%;">Подпись</td> <td style="width: 15%;">Дата</td> </tr> <tr> <td> </td> <td> </td> <td> </td> <td> </td> <td> </td> </tr> </table> </div> <div style="width: 50%; text-align: center;"> <p>ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)</p> </div> <div style="width: 10%; text-align: right;"> <p>Лист</p> <p>48</p> </div> </div>					Изм.	Лист	№ докум.	Подпись	Дата																																						
Изм.		Лист	№ докум.	Подпись	Дата																																												

Перв. примен.	<p>программы представлены на рисунке 3.2 (синие прямоугольники – позиции объектов, записанные в файлах описания изображения, красные прямоугольники – результаты локализации объектов):</p>				
	<div>   </div>				
Справ. №					
Подпись и дата					
Инв. № дубл.					
Взам. инв. №					
Подпись и дата					
Инв. № подл.					
<p>Рисунок 3.2. Результаты работы программного продукта по локализации объектов на изображении</p>					Лист
<div> <div>ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)</div> <div>49</div> </div>					
Изм.	Лист	№ докум.	Подпись	Дата	

Перв. примен.	3.3. Анализ результатов эксперимента																																																																																																																																																												
	Результаты тестирования алгоритма в нормальных условиях приведены в таблице 3.7:																																																																																																																																																												
Справ. №	Таблица 3.7. Результаты тестирования алгоритма в нормальных условиях																																																																																																																																																												
	<table><tr><th>Коэффициент отбора k</th><th>Размер области n, пикс</th><th>Смещение m, пикс</th><th>Метод кластеризации</th><th>Точность локализации, %</th><th>Ошибки 1 рода, %</th><th>Ошибки 2 рода, %</th><th>Время обработки, с</th></tr><tr><td rowspan="12">0,6</td><td rowspan="4">16</td><td rowspan="2">4</td><td>K-means</td><td>95,50</td><td>1,00</td><td>3,50</td><td>0,14</td></tr><tr><td>Mean shift</td><td>96,50</td><td>1,00</td><td>2,50</td><td>0,14</td></tr><tr><td rowspan="2">8</td><td>K-means</td><td>82,00</td><td>7,00</td><td>11,00</td><td>0,03</td></tr><tr><td>Mean shift</td><td>83,50</td><td>9,00</td><td>7,50</td><td>0,03</td></tr><tr><td rowspan="4">32</td><td rowspan="2">4</td><td>K-means</td><td>96,50</td><td>1,00</td><td>2,50</td><td>0,22</td></tr><tr><td>Mean shift</td><td>96,50</td><td>1,00</td><td>2,50</td><td>0,24</td></tr><tr><td rowspan="2">8</td><td>K-means</td><td>88,00</td><td>5,00</td><td>7,00</td><td>0,03</td></tr><tr><td>Mean shift</td><td>90,50</td><td>6,00</td><td>3,50</td><td>0,03</td></tr><tr><td rowspan="4">64</td><td rowspan="2">4</td><td>K-means</td><td>93,00</td><td>3,00</td><td>4,00</td><td>0,39</td></tr><tr><td>Mean shift</td><td>92,50</td><td>3,00</td><td>4,50</td><td>0,48</td></tr><tr><td rowspan="2">8</td><td>K-means</td><td>84,00</td><td>11,00</td><td>5,00</td><td>0,05</td></tr><tr><td>Mean shift</td><td>85,00</td><td>11,00</td><td>4,00</td><td>0,05</td></tr><tr><td rowspan="12">0,65</td><td rowspan="4">16</td><td rowspan="2">4</td><td>K-means</td><td>92,00</td><td>3,00</td><td>5,00</td><td>0,12</td></tr><tr><td>Mean shift</td><td>92,00</td><td>3,00</td><td>5,00</td><td>0,13</td></tr><tr><td rowspan="2">8</td><td>K-means</td><td>75,00</td><td>15,50</td><td>9,50</td><td>0,03</td></tr><tr><td>Mean shift</td><td>73,00</td><td>19,50</td><td>7,50</td><td>0,03</td></tr><tr><td rowspan="4">32</td><td rowspan="2">4</td><td>K-means</td><td>94,00</td><td>2,50</td><td>3,50</td><td>0,17</td></tr><tr><td>Mean shift</td><td>95,00</td><td>2,50</td><td>2,50</td><td>0,19</td></tr><tr><td rowspan="2">8</td><td>K-means</td><td>84,50</td><td>10,50</td><td>5,00</td><td>0,03</td></tr><tr><td>Mean shift</td><td>84,50</td><td>11,50</td><td>4,00</td><td>0,03</td></tr><tr><td rowspan="4">64</td><td rowspan="2">4</td><td>K-means</td><td>90,50</td><td>6,00</td><td>3,50</td><td>0,28</td></tr><tr><td>Mean shift</td><td>89,00</td><td>6,00</td><td>5,00</td><td>0,33</td></tr><tr><td rowspan="2">8</td><td>K-means</td><td>79,00</td><td>14,00</td><td>6,50</td><td>0,04</td></tr><tr><td>Mean shift</td><td>78,50</td><td>14,00</td><td>7,50</td><td>0,04</td></tr></table>										Коэффициент отбора k	Размер области n, пикс	Смещение m, пикс	Метод кластеризации	Точность локализации, %	Ошибки 1 рода, %	Ошибки 2 рода, %	Время обработки, с	0,6	16	4	K-means	95,50	1,00	3,50	0,14	Mean shift	96,50	1,00	2,50	0,14	8	K-means	82,00	7,00	11,00	0,03	Mean shift	83,50	9,00	7,50	0,03	32	4	K-means	96,50	1,00	2,50	0,22	Mean shift	96,50	1,00	2,50	0,24	8	K-means	88,00	5,00	7,00	0,03	Mean shift	90,50	6,00	3,50	0,03	64	4	K-means	93,00	3,00	4,00	0,39	Mean shift	92,50	3,00	4,50	0,48	8	K-means	84,00	11,00	5,00	0,05	Mean shift	85,00	11,00	4,00	0,05	0,65	16	4	K-means	92,00	3,00	5,00	0,12	Mean shift	92,00	3,00	5,00	0,13	8	K-means	75,00	15,50	9,50	0,03	Mean shift	73,00	19,50	7,50	0,03	32	4	K-means	94,00	2,50	3,50	0,17	Mean shift	95,00	2,50	2,50	0,19	8	K-means	84,50	10,50	5,00	0,03	Mean shift	84,50	11,50	4,00	0,03	64	4	K-means	90,50	6,00	3,50	0,28	Mean shift	89,00	6,00	5,00	0,33	8	K-means	79,00	14,00	6,50	0,04	Mean shift	78,50	14,00	7,50
Коэффициент отбора k	Размер области n, пикс	Смещение m, пикс	Метод кластеризации	Точность локализации, %	Ошибки 1 рода, %	Ошибки 2 рода, %	Время обработки, с																																																																																																																																																						
0,6	16	4	K-means	95,50	1,00	3,50	0,14																																																																																																																																																						
			Mean shift	96,50	1,00	2,50	0,14																																																																																																																																																						
		8	K-means	82,00	7,00	11,00	0,03																																																																																																																																																						
			Mean shift	83,50	9,00	7,50	0,03																																																																																																																																																						
	32	4	K-means	96,50	1,00	2,50	0,22																																																																																																																																																						
			Mean shift	96,50	1,00	2,50	0,24																																																																																																																																																						
		8	K-means	88,00	5,00	7,00	0,03																																																																																																																																																						
			Mean shift	90,50	6,00	3,50	0,03																																																																																																																																																						
	64	4	K-means	93,00	3,00	4,00	0,39																																																																																																																																																						
			Mean shift	92,50	3,00	4,50	0,48																																																																																																																																																						
		8	K-means	84,00	11,00	5,00	0,05																																																																																																																																																						
			Mean shift	85,00	11,00	4,00	0,05																																																																																																																																																						
0,65	16	4	K-means	92,00	3,00	5,00	0,12																																																																																																																																																						
			Mean shift	92,00	3,00	5,00	0,13																																																																																																																																																						
		8	K-means	75,00	15,50	9,50	0,03																																																																																																																																																						
			Mean shift	73,00	19,50	7,50	0,03																																																																																																																																																						
	32	4	K-means	94,00	2,50	3,50	0,17																																																																																																																																																						
			Mean shift	95,00	2,50	2,50	0,19																																																																																																																																																						
		8	K-means	84,50	10,50	5,00	0,03																																																																																																																																																						
			Mean shift	84,50	11,50	4,00	0,03																																																																																																																																																						
	64	4	K-means	90,50	6,00	3,50	0,28																																																																																																																																																						
			Mean shift	89,00	6,00	5,00	0,33																																																																																																																																																						
		8	K-means	79,00	14,00	6,50	0,04																																																																																																																																																						
			Mean shift	78,50	14,00	7,50	0,04																																																																																																																																																						
Подпись и дата	Инов. № дубл.	Взам. инв. №	Инов. № подл.	Подпись и дата																																																																																																																																																									
					ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)																																																																																																																																																								
					Лист 50																																																																																																																																																								
Изм.	Лист	№ докум.	Подпись	Дата																																																																																																																																																									

Перв. примен.	Таблица 3.7. Результаты тестирования алгоритма в нормальных условиях							
	Кoeffициент отбора k	Размер области n, пикс	Смещение m, пикс	Метод кластеризации	Точность локализации, %	Ошибки 1 рода, %	Ошибки 2 рода, %	Время обработки, с
0,7	16	4	K-means	82,50	7,00	10,50	0,11	
			Mean shift	85,00	7,00	8,00	0,11	
		8	K-means	60,00	29,00	11,00	0,03	
			Mean shift	52,00	37,00	11,00	0,03	
	32	4	K-means	89,00	6,50	4,50	0,14	
			Mean shift	89,00	7,00	4,00	0,14	
		8	K-means	80,00	13,50	6,50	0,03	
			Mean shift	80,00	15,00	5,00	0,03	
	64	4	K-means	82,00	10,00	8,00	0,21	
			Mean shift	80,50	10,00	9,50	0,23	
		8	K-means	73,50	18,50	8,00	0,03	
			Mean shift	71,50	18,50	10,00	0,03	

Таким образом, наилучшая точность локализации объекта на изображении (96,5 %) была достигнута при следующих конфигурациях запуска алгоритма:

1. Коэффициент отбора ключевых точек k = 0,6, размер области n = 16, шаг разбиения на области m = 4, метод кластеризации – Mean shift. Среднее время обработки одного изображения – 7,97 с.
2. Коэффициент отбора ключевых точек k = 0,6, размер области n = 32, шаг разбиения на области m = 4, метод кластеризации – k-means. Среднее время обработки одного изображения – 11,84 с.
3. Коэффициент отбора ключевых точек k = 0,6, размер области n = 32, шаг разбиения на области m = 4, метод кластеризации – Mean shift. Среднее время обработки одного изображения – 13,11 с.

Учитывая, что в используемой базе изображений для каждого объекта имеются в том числе и изображения с отличными от шаблона масштабом и углом поворота объекта, можно сделать вывод о том, что разработанный алгоритм является устойчивым к таким изменениям объекта. Это объясняется тем, что предложенный алгоритм не учитывает взаимное расположение ключевых точек на объекте, а на исследуемом изображении ищет область с максимальным скоплением точек, эквивалентных ключевым. Шаблон объекта независимо от угла поворота содержит одинаковое количество ключевых точек, именно поэтому объект может быть

Подпись и дата	Инов. № дубл.	Взам. инв. №	Подпись и дата	Инов. № подл.

Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)	Лист 51
------	------	----------	---------	------	---	------------

Перв. примен.					
Справ. №					
Подпись и дата					
Инв. № дубл.					
Взам. инв. №					
Подпись и дата					
Инв. № подл.					

локализован на изображении с одинаковой точностью при разных углах поворота. Аналогичное утверждение справедливо и для шаблонов, имеющих отличный от исследуемого изображения масштаб объекта, за исключением случаев, когда размер объекта на шаблоне настолько мал, что не позволяет обнаруживать достаточное количество ключевых точек.

Все три полученные конфигурации были использованы для дальнейшего тестирования разработанного метода в условиях наличия шумов различной интенсивности. Результаты данного тестирования приведены в таблице 3.8:

Таблица 3.8. Результаты тестирования алгоритма в условиях наличия шумов

Коэффициент k	Размер области n, пикс	Смещение m, пикс	Метод кластеризации	ОС Ш, дБ	Точность локализации, %	Ошибки 1 рода, %	Ошибки 2 рода, %
0,6	16	4	Mean shift	0	54,50	45,50	0,00
				10	87,50	12,50	0,00
				20	94,50	5,50	0,00
0,6	32	4	K-means	0	74,00	26,00	0,00
				10	92,00	8,00	0,00
				20	94,00	6,00	0,00
0,6	32	4	Mean shift	0	79,00	21,00	0,00
				10	90,50	9,50	0,00
				20	95,00	4,50	0,50

Таким образом, наибольшая точность локализации объектов на изображении при значениях ОСШ 0 дБ (79 %) и 20 дБ (95 %) была достигнута при конфигурации запуска со следующими параметрами: коэффициент отбора ключевых точек $k = 0,6$, размер области $n = 32$, шаг разбиения на области $m = 4$, метод кластеризации – Mean shift.

При значении ОСШ 10 дБ наивысшая точность локализации (92 %) была достигнута при конфигурации алгоритма: коэффициент отбора ключевых точек $k = 0,6$, размер области $n = 32$, шаг разбиения на области $m = 4$, метод кластеризации – k-means.

Также стоит отметить значительное превосходство ошибок 1-го рода (ложноположительных срабатываний) над ошибками 2-го рода (ложноотрицательными срабатываниями) при работе алгоритма в условиях наличия шумов.

В таблице 3.9 представлены результаты работы известных методов локализации объектов на изображении, рассмотренных в главе 1.

Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)	Лист
						52

Перв. примен.	Таблица 3.9. Результаты работы известных методов локализации				
	Способ формирования признакового описания объекта	Способ принятия решения о локализации объекта	Точность локализации, %		
Справ. №	Инвариантное к масштабу признаковое преобразование (SIFT)	Локализация объекта на основе объединения регионов изображения по принципу схожести	83,33 – 91,66		
	Ускоренное построение устойчивых признаков (SURF)	Иерархическая пирамидальная схема локализации объекта	95		
	Двоичные устойчивые независимые элементарные признаки (BRIEF)	Классификатор в виде вероятностных деревьев	94 – 98		
<p>Сравнивая данные из таблицы 3.9 и полученные результаты тестирования алгоритма в различных условиях (таблица 3.7 и таблица 3.8), можно сделать вывод о том, что разработанный метод имеет точность локализации объектов на изображении, вполне сопоставимую с точностью рассмотренных известных методов. По сравнению с некоторыми известными методами точность предложенного подхода к решению задачи оказывается выше.</p>					
Подпись и дата					
Инв. № дубл.					
Взам. инв. №					
Подпись и дата					
Инв. № подл.					
<div style="display: flex; justify-content: space-between;"> Изм. Лист № докум. Подпись Дата </div>					Лист
ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)					53

Перв. примен.						<h3 style="text-align: center;">3.4. Выводы по главе 3</h3> <p>Данная глава посвящена описанию и проведению вычислительного эксперимента по исследованию предлагаемого метода решения задачи локализации объектов на изображении, а также анализу результатов этого эксперимента. Помимо этого, в данной главе описывается база изображений, собранная для проведения вычислительного эксперимента, и программный продукт, реализующий разработанный алгоритм решения задачи.</p> <p>В результате проведения тестирования алгоритма в нормальных условиях и в условиях наложения шумов были выявлены конфигурации запуска, дающие наибольшую точность локализации, и конфигурации, наиболее устойчивые к появлению шумов различной интенсивности. Были получены следующие значения точности локализации объектов на изображении:</p> <ul style="list-style-type: none"> • При отсутствии шума: 96,5 %. • При ОСШ 20 дБ: 95 %. • При ОСШ 10 дБ: 92 %. • При ОСШ 0 дБ: 79 %. <p>Данные результаты говорят о том, что предлагаемый метод решения задачи является устойчивым к появлению шума на изображении искомого объекта.</p> <p>Также полученные результаты были сравнены с результатами работы некоторых известных методов локализации объектов на изображении. В ходе этого сравнения было выявлено, что предлагаемый метод имеет вполне сопоставимую с результатами работы существующих методов точность локализации. Это говорит о том, что разработанный алгоритм является вполне конкурентоспособным по сравнению с аналогами.</p>				
	Справ. №									
Подпись и дата										
	Инв. № дубл.									
Взам. инв. №										
	Подпись и дата									
Инв. № подл.										
	Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)				
					54					

Перв. примен.	<h2 style="text-align: center;">Заключение</h2> <p>В ходе выполнения выпускной квалификационной работы был составлен обзор существующих известных методов решения задачи локализации объектов на изображении. На основании этого обзора были сформулированы основные проблемы существующих методов, возникающие на различных этапах решения задачи.</p> <p>С учётом информации, полученной в результате составления обзора был предложен новый метод решения задачи локализации объекта на изображении. Данный метод заключается в использовании новой комбинации подходов к решению рассматриваемой задачи на различных этапах. На этапе формирования признакового описания объекта используются методы, основанные на теории активного восприятия. На этапе принятия решения о локализации объекта – методы, основанные на кластеризации особых точек объекта с помощью алгоритмов k-means и Mean shift.</p> <p>Для реализации предложенного метода был разработан программный продукт на языке R. Разработанный алгоритм был протестирован на различных наборах входных данных в ходе проведенного вычислительного эксперимента. Полученные при проведении эксперимента результаты свидетельствуют о корректной работе предложенного метода при решении задачи локализации объекта на изображении.</p> <p>Величина точности локализации объекта при использовании разработанного метода оказалась вполне сопоставимой с точностью рассмотренных в обзоре методов, а в некоторых случаях даже выше. Это говорит о том, что данный метод является конкурентоспособным и пригодным к использованию на практике.</p> <p>Исходя из указанного выше можно сделать вывод о том, что поставленные перед началом научного исследования задачи были выполнены в полном объёме, а цель работы – достигнута.</p>					
						Справ. №
Подпись и дата	Инв. № дубл.	Взам. инв. №	Подпись и дата	Инв. № подл.		
Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)	Лист
						55

Перв. примен.					
Справ. №					
<p>17. Утробин В.А. Компьютерная обработка изображений. Информационные модели этапа понимания. – Нижний Новгород: НГТУ им. Р.Е. Алексеева, 2006. – С. 247.</p> <p>18. Утробин В.А. Информационные модели системы зрительного восприятия для задач компьютерной обработки изображений. – Нижний Новгород: НГТУ им. Р.Е. Алексеева, 2001. – С. 234.</p> <p>19. Утробин В.А. Физические интерпретации элементов алгебры изображения // Успехи физических наук. – 2004. – Т. 174, № 10. – С. 1089-1104.</p> <p>20. Leemis L. Learning Base R. – Lightning Source. – 2016.</p> <p>21. Faster Gibbs sampling MCMC from within R – Darren Wilkinson’s research blog. – URL: https://darrenjw.wordpress.com/2011/07/31/faster-gibbs-sampling-mcmc-from-within-r/</p>					
Подпись и дата					
Инв. № дубл.					
Взам. инв. №					
Подпись и дата					
Инв. № подл.					
Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М15-ИВТ-3)-021-2017 (ПЗ)
					Лист 57