

Министерство образования и науки Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**



Институт кибернетики
Направление подготовки 09.04.01 Информатика и вычислительная техника
Кафедра программной инженерии

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Тема работы
Программное обеспечение анализа изменения ландшафтного покрова по данным камеры высокого разрешения, управляемой с БПЛА с фиксированным крылом УДК 004.42:504.54:528.71:629.73-519

Студент

Группа	ФИО	Подпись	Дата
8ВМ5В	Б.С.Бокарев		

Руководитель

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент каф. ПИ	С.В. Аксёнов	К.Т.Н.		

КОНСУЛЬТАНТЫ:

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент каф. МЕН	В.Ю. Конотопский	К.Э.Н.		

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
ассистент каф. ЭБЖ	П.А. Акулов			

ДОПУСТИТЬ К ЗАЩИТЕ:

Зав. кафедрой	ФИО	Ученая степень, звание	Подпись	Дата
ПИ	М.А. Иванов	К.Т.Н.		

Планируемые результаты обучения по ООП

Код результата	Результат обучения (выпускник должен быть готов)
<i>Общепрофессиональные компетенции</i>	
P1	Воспринимать и самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте.
P2	Владеть и применять методы и средства получения, хранения, переработки и трансляции информации посредством современных компьютерных технологий, в том числе в глобальных компьютерных сетях.
P3	Демонстрировать культуру мышления, способность выстраивать логику рассуждений и высказываний, основанных на интерпретации данных, интегрированных из разных областей науки и техники, выносить суждения на основании неполных данных, анализировать профессиональную информацию, выделять в ней главное, структурировать, оформлять и представлять в виде аналитических обзоров с обоснованными выводами и рекомендациями.
P4	Анализировать и оценивать уровни своих компетенций в сочетании со способностью и готовностью к саморегулированию дальнейшего образования и профессиональной мобильности. Владеть, по крайней мере, одним из иностранных языков на уровне социального и профессионального общения, применять специальную лексику и профессиональную терминологию языка.
<i>Профессиональные компетенции</i>	
P5	Выполнять инновационные инженерные проекты по разработке аппаратных и программных средств автоматизированных систем различного назначения с использованием современных методов проектирования, систем автоматизированного проектирования, передового опыта разработки конкурентно способных изделий.
P6	Планировать и проводить теоретические и экспериментальные исследования в области проектирования аппаратных и программных средств автоматизированных систем с использованием новейших достижений науки и техники, передового отечественного и зарубежного опыта. Критически оценивать полученные данные и делать выводы.
P7	Осуществлять авторское сопровождение процессов проектирования, внедрения и эксплуатации аппаратных и программных средств автоматизированных систем различного назначения.
<i>Общекультурные компетенции</i>	
P8	Использовать на практике умения и навыки в организации исследовательских, проектных работ и профессиональной эксплуатации современного оборудования и приборов, в управлении коллективом.
P9	Осуществлять коммуникации в профессиональной среде и в обществе в

	целом, активно владеть иностранным языком, разрабатывать документацию, презентовать и защищать результаты инновационной инженерной деятельности, в том числе на иностранном языке
P10	Совершенствовать и развивать свой интеллектуальный и общекультурный уровень. Проявлять инициативу, в том числе в ситуациях риска, брать на себя всю полноту ответственности.
P11	Демонстрировать способность к самостоятельному обучению новым методам исследования, к изменению научного и научно-производственного профиля своей профессиональной деятельности, способность самостоятельно приобретать с помощью информационных технологий и использовать в практической деятельности новые знания и умения, в том числе в новых областях знаний, непосредственно не связанных со сферой деятельности, способность к педагогической деятельности.

Министерство образования и науки Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**



Институт кибернетики
Направление подготовки 09.04.01 Информатика и вычислительная техника
Кафедра программной инженерии

УТВЕРЖДАЮ:
Зав. кафедрой
_____ М.А. Иванов
_____ 2017 г.

ЗАДАНИЕ
на выполнение выпускной квалификационной работы

В форме:

Магистерской диссертации

Студенту:

Группа	ФИО
8ВМ5В	Бокарев Борис Сергеевич

Тема работы:

Программное обеспечение анализа изменения ландшафтного покрова по данным камеры высокого разрешения, управляемой с БПЛА с фиксированным крылом

Утверждена приказом директора

№1652/с от 13.03.2017

Срок сдачи студентом выполненной работы:

08.06.2017 г.

ТЕХНИЧЕСКОЕ ЗАДАНИЕ:

Исходные данные к работе

наименование объекта исследования – алгоритм поиска текстуры на основе сравнения энергетических характеристик Лавса для образца и обрабатываемого кадра;
требования к продукту – программа должна обеспечивать поиск текстур в изображениях. В результате исследования необходимо предложить модификацию метода. Программа должна отвечать требованиям качества обработки и производительности.

Перечень подлежащих исследованию, проектированию и разработке вопросов	<i>аналитический обзор по литературным источникам с целью выяснения достижений мировой науки техники в рассматриваемой области; постановка задачи исследования; содержание процедуры исследования; обсуждение результатов выполненной работы; заключение по работе.</i>
Перечень графического материала	<i>примеры работы реализованной программы</i>

Консультанты по разделам выпускной квалификационной работы

Раздел	Консультант
Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	<i>В.Ю. Конотопский, к.э.н., доцент каф. МЕН</i>
Социальная ответственность	<i>П.А. Акулов, ассистент каф. ЭБЖ</i>
Раздел, выполненный на иностранном языке	<i>Т.В. Сидоренко, зав. каф. ИЯ ИК</i> <i>Е.С. Чердынцев, к.т.н., доцент каф. ПИ</i>

Названия разделов, которые должны быть написаны на русском и иностранном языках:

1. Способы описания содержания изображения
2. Реализация метода поиска текстур
3. Оценка результатов работы предлагаемого метода
4. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение
5. Социальная ответственность
6. Video stream segmentation

Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику	
---	--

Задание выдал руководитель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент каф. ПИ	С.В. Аксёнов	к.т.н.		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8ВМ5В	Б.С. Бокарев		

Министерство образования и науки Российской Федерации
 Федеральное государственное автономное образовательное учреждение
 высшего образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
 ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**



Институт кибернетики
 Направление подготовки 09.04.01 Информатика и вычислительная техника
 Уровень образования магистр
 Кафедра программной инженерии
 Период выполнения _____ весенний семестр 2016/2017 учебного года

Форма представления работы:

Магистерская диссертация

**КАЛЕНДАРНЫЙ РЕЙТИНГ-ПЛАН
 выполнения выпускной квалификационной работы**

Срок сдачи студентом выполненной работы:	08.06.2017
--	------------

Дата контроля	Название раздела (модуля) / вид работы (исследования)	Максимальный балл раздела (модуля)
14.04.2017	1. Способы описания содержания изображения	10
22.04.2017	2. Реализация метода поиска текстур	20
05.05.2017	3. Оценка результатов работы предлагаемого метода	30
16.05.2017	4. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	15
23.05.2017	5. Социальная ответственность	15
30.05.2017	6. Video stream segmentation	10

Составил преподаватель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент каф. ПИ	С.В. Аксёнов	к.т.н.		

СОГЛАСОВАНО:

Зав. кафедрой	ФИО	Ученая степень, звание	Подпись	Дата
Доцент каф. ПИ	М.А. Иванов	к.т.н.		

Реферат

Выпускная квалификационная работа содержит 110 страниц, 24 рисунка, 17 таблиц, 39 источников, 1 приложение.

К ключевым словам, используемым в данном документе, относятся: видеообработка, текстурный анализ, текстурные признаки, энергетические характеристики Лавса, поиск в изображениях по образцу.

Объектом исследования является метод поиска текстуры на основе сравнения энергетических характеристик Лавса для образца и обрабатываемого кадра. Цель работы – разработка программного обеспечения для поиска текстур в изображениях.

В процессе исследования была изучена предметная область проблемы, существующие решения, реализовано и протестировано приложение, рассчитаны показатели безопасности жизнедеятельности и проведен анализ возможной экономической эффективности. Программа реализована с использованием языка программирования C#, среды разработки Microsoft Visual Studio 2010 и Emgu CV – кроссплатформенной обертки графической библиотеки OpenCV. В результате было разработано приложение, получающее на вход цветные или полутоновые изображения текстуры, размер окрестности для расчета энергетических карт Лавса и видеофайл. На выходе получаем ключевые кадры видеофайла, содержащие образцы текстур.

Анализ результатов выполнения программы для различных параметров показал, что точность поиска зависит от соответствия размеров образцов текстуры и исходного кадра. Необходимо экспериментальным путем определять параметры для отдельных случаев.

Разработанная программа может применяться в системах распознавания видеофайлов, для организации поиска в коллекциях видео и изображений. В настоящее время программа находится на стадии дальнейшего улучшения, задачей которого является повышение

производительности и качества поиска путем модификации метода выделения характерных признаков.

Определения

В данной работе применены следующие термины с соответствующими определениями.

видеопоток: Временная последовательность кадров определенного формата, закодированная в битовый поток.

кадр: Элемент цифрового видеопотока, отдельное изображение.

ключевой кадр: Отдельный кадр, несущий полную информацию о содержании видефрагмента.

текстура: Понятие, характеризующее и измеряющее параметры однородности изображения.

энергетические характеристики Лавса: Характеристики, описывающие свойства типа ребер, скачков, волн на изображении.

энергетические карты Лавса: совокупность энергетические характеристики Лавса, организованная в виде двумерной матрицы размерностью, совпадающей с обрабатываемым изображением.

сегментация: Разбиение изображения на множество покрывающих его областей.

гистограмма: это график статистического распределения элементов цифрового изображения с различной яркостью, в котором по горизонтальной оси представлена яркость, а по вертикали – относительное число пикселей с конкретным значением яркости.

Обозначения и сокращения

ПЭВМ – персональные электронная вычислительная машина;

ВДТ – видеодисплейный терминал;

КЕО – коэффициент естественной освещенности;

НТИ – научно-техническое исследование.

Содержание

Введение.....	15
1 Способы описания содержания изображения.....	17
1.1 Цветовые признаки	17
1.1.1 Построение цветового квадратомиического дерева.....	17
1.1.2 Сопоставление гистограмм по тону и интенсивности	19
1.2 Поиск по форме.....	25
1.2.1 Пороговая обработка	26
1.2.2 Марковская фильтрация.....	29
1.2.3 Граничные алгоритмы	31
1.2.4 Алгоритм разрастания регионов	32
1.3 Текстурные характеристики	35
1.3.1 Структурный способ представления текстур.....	35
1.3.2 Статистический подход к описанию текстур.....	39
2 Реализация метода поиска текстур в видеопотоке	49
2.1 Выбор и обоснование среды разработки	49
2.2 Библиотека машинного зрения Emgu CV.....	50
2.3 Сравнение энергетических карт	51
2.4 Комбинация текстурных и цветовых признаков	53
2.5 Реализация обработки видеоданных	54
3 Оценка результатов работы предлагаемого метода	56
3.1 Влияние масштаба изображения на эффективность	56
3.2 Влияние количества образцов текстур на результат	58
3.4 Подбор оптимальных параметров	60

4	Финансовый менеджмент, ресурсоэффективность и ресурсосбережение...	63
4.1	Организация и планирование работ	63
4.1.1	Продолжительность этапов работ	64
4.1.2	Расчет накопления готовности проекта.....	68
4.2	Расчет сметы затрат на выполнение проекта	69
4.2.1	Расчет затрат на материалы	70
4.2.2	Расчет заработной платы.....	70
4.2.3	Расчет затрат на социальный налог	72
4.2.4	Расчет затрат на электроэнергию	72
4.2.5	Расчет амортизационных расходов	73
4.2.6	Расчет расходов на услуги связи	74
4.2.7	Расчет прочих расходов.....	74
4.2.8	Расчет общей себестоимости разработки.....	75
4.3	Оценка научно-технического уровня НИР.....	76
4.4	Экономический эффект НИР	77
5.	Социальная ответственность	80
5.1	Введение.....	80
5.2	Описание рабочего места	81
5.3	Производственная и эксплуатационная безопасность	82
5.3.1	Электробезопасность	82
5.3.2	Микроклимат	84
5.3.3	Освещенность.....	85
5.3.4	Монотонный режим работы.....	86
5.3.5	Пожарная безопасность.....	88

5.4 Рекомендации по минимизации влияния вредных и опасных факторов производства	88
5.4.1 Рекомендации по защите от электрического тока	88
5.4.2 Рекомендации по улучшению микроклимата	89
5.4.3 Рекомендации по минимизации влияния освещения	89
5.4.4 Меры по обеспечению пожарной безопасности.....	90
5.5 Экологическая безопасность.....	90
5.5.1 Анализ воздействия на окружающую среду	90
5.5.2 Рекомендации по минимизации влияния на окружающую среду	90
5.6 Безопасность в чрезвычайных ситуациях.....	91
5.6.1 Перечень возможных чрезвычайных ситуаций на объекте.....	91
5.6.2 Меры по ликвидации чрезвычайных ситуаций и последствий.....	92
5.7 Правовые и организационные вопросы обеспечения безопасности ...	92
5.7.1 Психофизиологические факторы	92
5.7.2 Организационные мероприятия при компоновке рабочей зоны ..	93
Заключение	94
Список используемых источников.....	96
6 Video stream segmentation	102
6.1 Pixel-based methods.....	102
6.2 Histogram analysis methods	104
6.3 Block-based methods	105
6.4 Feature extraction.....	107
6.5 Motion-based methods	108
6.6 Video keyframe extraction.....	110

Conclusion..... 111

Введение

Исследованию алгоритмов обработки видеоданных в настоящее время посвящено множество работ. Особое внимание уделяется методам описания содержимого видеопотока, поскольку с их помощью можно решить такие актуальные задачи как сегментация, выделение ключевых кадров, индексация, поиск видео по образцу и распознавание объектов в видеопотоке.

Целью данной работы является разработка программного обеспечения поиска заданных текстур в видеопотоке. Текстуры характеризуют объекты окружающего мира и являются естественными для восприятия человеческим глазом. Они объединяют в себе характеристики формы, взаимного расположения элементов, а также цвета.

Объектом данного исследования является метод выделения и сравнения энергетических текстурных характеристик Лавса. В работе проводится оценка его точности и ее зависимость от различных параметров, таких как разрешение исходного кадра, настраиваемый размер окрестности, количество образцов текстур.

Для обработки цветных текстур предлагается использовать объединенные признаки текстуры и цвета, что позволяет более подробно описать содержимое видеопотока и повысить качество поиска.

Задачи обработки графических данных с выделенными текстурами возникают при анализе спутниковых снимков для распознавания различных областей, например, пожары, города, леса, водоемы и т.д. В медицине широко применяются алгоритмы поиска патологических нарушений, объемов тканей на рентгеновских снимках. Предложенный метод может быть применен для обработки видеозаписей наружных веб-камер.

Другим направлением является поиск видео и изображений по образцу для организации доступа к графической информации в базах данных.

Подобные задачи характерны для поисковых систем, интернет-магазинов, а также частных коллекций видеофайлов.

1 Способы описания содержания изображения

В настоящее время для поиска изображений по содержанию используется подход, основанный на сравнении характерных признаков. Задачей поиска изображений по содержанию (англ. Content-based image retrieval, CBIR) является выборка из заданной коллекции изображений, соответствующих определенному образцу. Одной из основных проблем эффективности подобных методов является описание признаков содержимого изображений таким образом, чтобы результат работы программы наименьшим образом зависел от ориентации, масштаба, наличия шума, теней и других аспектов, которые могут помешать однозначно описать изображение.

Именно по этой причине в современных алгоритмах широко используется сравнение цветовых составляющих. Наиболее распространенным методом является метод вычисления цветовых гистограмм, представленный в следующем разделе. Кроме того, в данной главе рассмотрены способы описания, основанные на текстурах и формах объектов, а также методы, используемые для обработки полутоновых изображений.

1.1 Цветовые признаки

1.1.1 Построение цветового квадратомиического дерева

Квадратомиическое дерево – это структура представления данных в двумерном пространстве. На первом уровне – изображение целиком, дерево имеет один узел. Следующий уровень состоит из четырех квадрантов, каждый из которого является новым узлом дерева. При очередном переходе к новому уровню каждый квадрант делится на четыре части до тех пор, пока на конечном уровне узел дерева не будет представлен единичным пикселем. На рисунке 1.1 изображены диаграмма квадратомиического дерева, включая

представление изображения для каждого уровня. Нулевой уровень – это корень дерева, который содержит информацию о цвете всех пикселей изображений. Первый уровень состоит из 4 квадрантов, второй – из 16 и т.д.

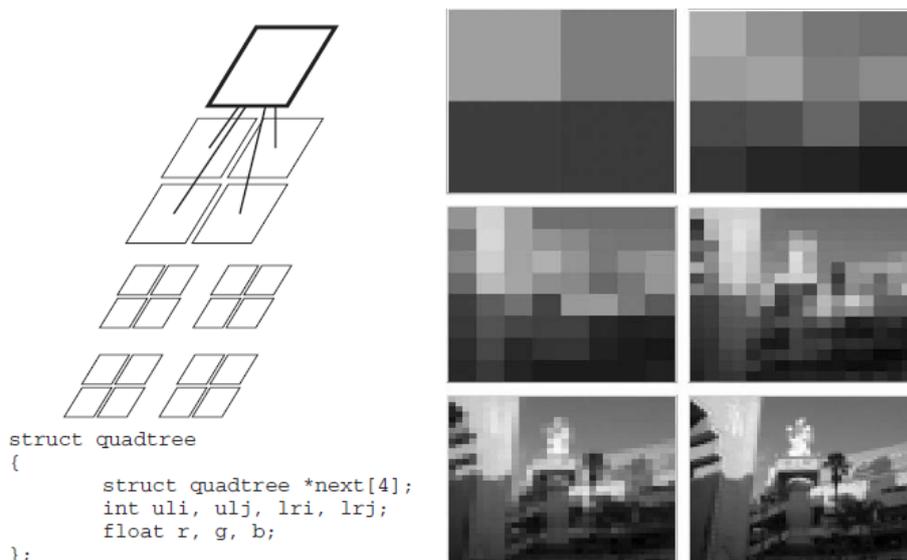


Рисунок 1.1 – Пример квадратомического дерева

На сегодняшний день существует множество путей сравнения изображений с помощью квадратомических деревьев. Например, можно определить уровень точности совпадения изображений относительно цветовых или некоторых других свойств. В некоторых частях изображения будет наблюдаться большее совпадение, что может также означать пространственное подобие. Однако более простым подходом будет построение вектора цветов, видимых на отдельном уровне, и использование его в качестве описательного признака.

При таком подходе корень дерева будет представлен одним цветом. Следующий уровень, состоящий из четырех элементов, каждый из которого характеризуется средним цветом, будет описан с помощью вектора из 4 цветов, или 12 компонент. Таким образом формируется признак квадрантов. Следующий признак позволяет описывать более точное подобие между изображениями и формируется путем представления вектора второго уровня квадратомического дерева, который состоит из 16 цветов, или 48 компонент.

Основным преимуществом использования квадротомических деревьев является их иерархическая структура, которая позволяет обеспечивать высокоэффективный доступ к элементам дерева. Поворот изображений, описанных с помощью квадротомических деревьев, на 90 градусов по и против часовой стрелки осуществляется простой заменой ориентации узлов. Для изменения разрешения объекта достаточно заменить все серые узлы второго уровня черными узлами [1].

Однако, несмотря на простоту манипулирования изображением, данный подход обладает рядом существенных недостатков, значительным образом затрудняющим применение предложенного метода для сравнения изображений. Поскольку при использовании квадротомических деревьев возможным является только поворот на 90 градусов, для повернутых изображений приходится заново перестраивать структуру. При этом сопоставление квадротомических деревьев исходного и повернутого изображений становится весьма сложной задачей [1].

1.1.2 Сопоставление гистограмм по тону и интенсивности

Использование квадротомических деревьев для выделения цветовых характеристик подразумевает связь между пространственными координатами и цветом. Гистограммы позволяют описывать непосредственно цветовую составляющую изображений вне зависимости от расположений объектов, что делает этот метод более точным.

Диапазон цветов представлен здесь прямоугольниками, ширину которых можно изменять самостоятельно. Поскольку в RGB-цветовом пространстве изображению соответствует трёхкоординатный вектор, данные характеристики редко используются при построении гистограмм. Чаще применяют скалярные величины, такие как тон или интенсивность.

Для построения гистограммы тона на практике сначала преобразуют изображение из привычного RGB- в HSV-цветовое пространство. В данной

модели яркостная информация (интенсивность) отделена от цветовой информации (цветовой тон, насыщенность). В результате модель HSV является идеальным средством для построения алгоритмов обработки изображений, поскольку в основе модели лежит естественное и интуитивно близкое человеку описание цвета [2].

Составляющая цветового тона H для каждого пикселя, представленного тремя характеристиками R , G и B , определяется по формуле (1.1) [3].

$$H = \begin{cases} \frac{(G - B)}{R + G - 2B}, & \text{если } B = \min(R, G, B) \\ \frac{(B - R)}{3(G + B - 2R)} + \frac{1}{3}, & \text{если } R = \min(R, G, B) \\ \frac{R - G}{3(R + B - 2G)} + \frac{2}{3}, & \text{если } G = \min(R, G, B) \end{cases} \quad (1.1)$$

При этом значение тона варьируется в пределах $0-360^\circ$. Однако в данном случае, характеристика должна принимать значение $0-180^\circ$. Алгоритм построения гистограммы сводится к созданию вектора нулей, состоящего из 180 элементов, каждый из которых соответствует определенному цветовому тону. Далее для каждого пикселя вычисляется значение H , которое должно укладываться в диапазон индексов вектора, и производится инкремент соответствующего элемента вектора.

Важно провести нормализацию гистограммы, в противном случае, сравнение изображений различных размеров будет невозможным. Нормализация включает в себя деление значения каждого элемента полученного вектора на их общую сумму, которая в большинстве случаев составляет число пикселей в изображении. Однако недостатком базового алгоритма является то, что при нем игнорируются точки с низким значением составляющей тона.

Две гистограммы H_1 и H_2 считаются равными, если выполняется выражение (1.2) для всех возможных значений i .

$$H_1[i] = H_2[i], \quad (1.2)$$

где $H_1[i], H_2[i]$ – i -ый элемент гистограммы.

Поскольку данное выражение будет выполняться только для абсолютно идентичных изображений, важно корректно определить допустимое отклонение. Кроме того, возникает сложность при сравнении с третьей гистограммой, когда необходимо выяснить, какой из образцов H_1 и H_2 наиболее близок к эталону.

Гистограммы представляют собой N -мерные массивы, которые могут рассматриваться в виде векторов, что делает возможным вычислить Евклидову метрику по формуле (1.3).

$$d(H_1, H_2) = \sqrt{\sum_i^N (H_1[i] - H_2[i])^2}, \quad (1.3)$$

где $d(H_1, H_2)$ – Евклидова метрика для гистограмм H_1 и H_2 ;

$H_1[i], H_2[i]$ – i -ый элемент гистограммы;

N – количество столбцов в гистограмме.

Подобным образом в большинстве программ определяется сходство гистограмм – как N -мерное расстояние между гистограммами, рассчитываемого поочередно для каждого столбца. Однако соседние столбцы гистограммы представляют близкие по тону точки изображения, в то время как при использовании Евклидовой метрики для обычного вектора смежные столбцы не имеют никакой связи.

В работе [3] для учета соседних столбцов предлагается использовать совокупные (кумулятивные) гистограммы, в которых значение столбца определяется суммой столбцов в обычной гистограмме для заданной точки – формула (1.4).

$$C[i] = C[i - 1] + H[i], \quad (1.4)$$

где $C[i]$ – значение i -го элемента кумулятивной гистограммы;

$H[i]$ – значение i -го элемента обычной гистограммы;

i – номер столбца, изменяется от 1 до N .

Для нормированной гистограммы, сумма среди всех столбцов равна 1. В кумулятивной гистограмме, значение элементов приближается к 1, т.к. значения столбцов заполняются в увеличивающемся порядке. Таким образом, различие между двумя гистограммами неоднократно суммируется среди одинаковых столбцов до тех пор, пока оно не станет равным. Это означает, что различие между двумя столбцами, лежащими далеко друг от друга, оказывает большее влияние, чем смежные столбцы.

Исследование эффективности предложенного метода по сравнению с классическим проводилось с использованием Амстердамской библиотеки изображений объектов (ALOI). Поиск, осуществлявшийся 1801 раз для разных образцов среди набора из 700000 изображений, в случае простой гистограммы показал успешный результат 65,9%. Для предложенного метода эффективность возросла на 3% и составила 69,1% положительных результатов [3].

В источнике [http://locv.ru/wiki/7.4.1_Сравнение_гистограмм] приведены другие способы сравнения гистограмм, реализованные в методах графической библиотеки OpenCV.

Например, по формуле (1.5) можно определить корреляцию, при расчете которой об абсолютном совпадении говорит значение 1, наихудшему результату соответствует минус 1, а при 0 считается, что гистограммы имеют случайные ассоциации (т.е. корреляция отсутствует).

$$d_{\text{correl}}(H_1, H_2) = \frac{\sum_i H'_1[i] \cdot H'_2[i]}{\sqrt{\sum_i H_1'^2[i] \cdot H_2'^2[i]}}, \quad (1.5)$$

где $d_{\text{correl}}(H_1, H_2)$ – значение корреляции для гистограмм H_1, H_2 ;

$H'_k[i]$ – значение, которое находится по формуле (1.6).

$$H'_k[i] = H_k[i] - \frac{\sum_j H_k[j]}{N}, \quad (1.6)$$

где N – число столбцов в гистограмме.

Другой способ метрики представлен в формуле (1.7) и называется хи-квадрат.

$$d_{\text{chi-square}}(H_1, H_2) = \sum_i \frac{(H_1[i] - H_2[i])^2}{H_1[i] + H_2[i]}, \quad (1.7)$$

где $d_{\text{chi-square}}(H_1, H_2)$ – хи-квадрат гистограмм H_1, H_2 .

В формуле (1.8) приведено выражение для расчета пересечения $d_{\text{intersection}}$.

$$d_{\text{intersection}}(H_1, H_2) = \sum_i \min(H_1[i], H_2[i]) \quad (1.8)$$

При наилучшем совпадении гистограмм значение пересечения будет приближаться к 1. Если пересечение равно 0, то гистограммы отличаются в значительной степени.

Совершенно иначе понимаются значения для соответствия Бхаттачария. В формуле (1.6) представлена формула для нахождения расстояния $d_{\text{Bhattacharya}}$.

$$d_{\text{Bhattacharya}}(H_1, H_2) = \sqrt{1 - \frac{\sum_i \sqrt{H_1[i] \cdot H_2[i]}}{\sum_i H_1[i] \cdot \sum_i H_2[i]}} \quad (1.9)$$

На рисунке 1.2 изображена сводная таблица метрик, приведенных в [5].

Histograms:		Matching measures:			
Model:		Correlation:	Chi square:	Intersection:	Bhattacharyya:
Exact match:		1.0	0.0	1.0	0.0
Half match:		0.7	0.67	0.5	0.55
Total mis-match:		-1.0	2.0	0.0	1.0

Рисунок 1.2 – Показатели совпадения гистограмм для корреляции, хи-квадрата, пересечения и расстояния Бхаттачария

Сравнительный анализ влияния условий фотосъемки на описанные выше метрики приведен в таблице 1.1 [6].

Таблица 1.1 – Сравнение гистограмм трех изображений

Сравнение	Корреляция	Хи-квадрат	Пересечение	Расстояния Бхаттачария
В помещении	0.96	0.14	0.82	0.2
В тени	0.09	1.57	0.13	0.8
На солнце	0.0	1.98	0.01	0.99

Сравнение проводилось для трех изображений ладони в различных условиях (рисунок 1.3).

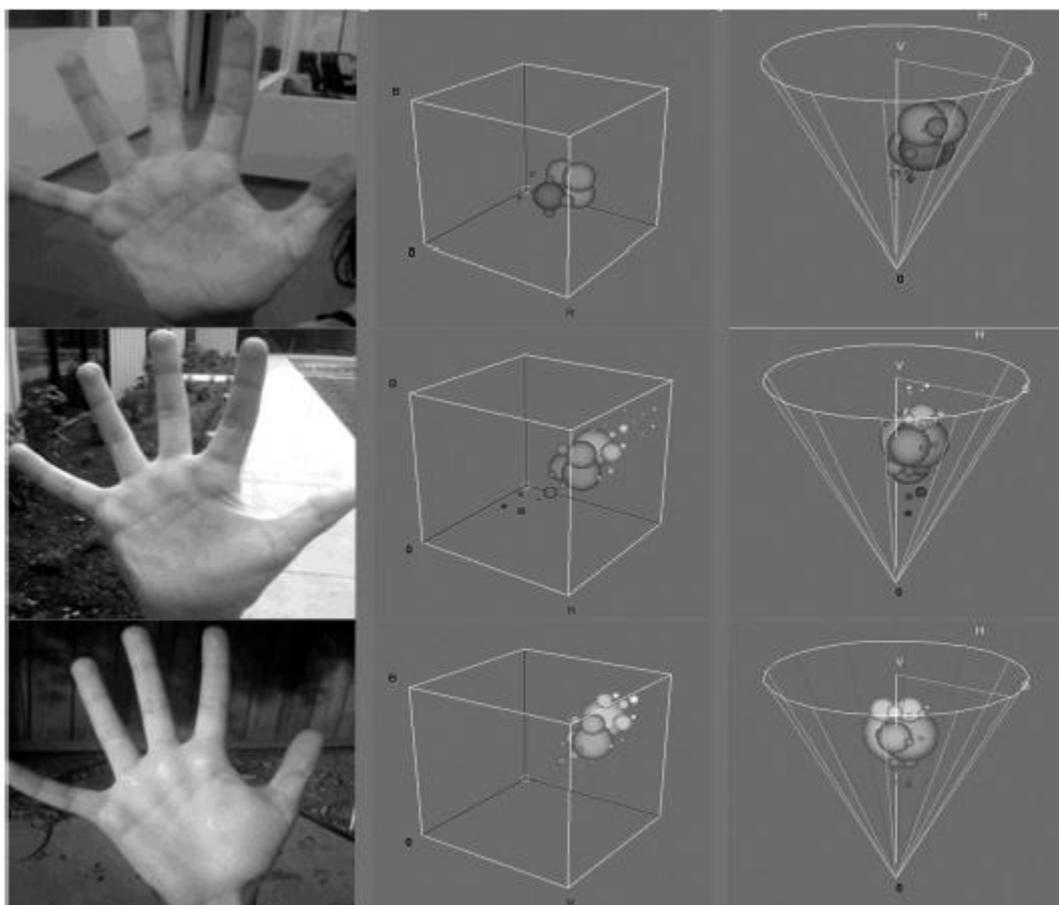


Рисунок 1.3 – Фотография для сравнения метрик гистограмм (посередине – гистограмма по каналам RGB, справа – в цветовом пространстве HSV)

При этом помимо использования цветового пространства HSV, в источнике также предлагается осуществлять сравнение в цветовом пространстве RGB, сопоставляя каждый канал по отдельности. Результаты расчета метрик в таблице 1.1 позволяют заметить, что условия освещения значительным образом влияют на выявленное различие между гистограммой для фотографий на рисунке 1.3 и эталонной, в качестве которой служила фотография ладони в помещении.

1.2 Поиск по форме

Другой подход к поиску по образцу заключается в сравнении форм некоторых областей изображений. Определение формы является задачей

сегментации или выделения границ на изображениях. Другие методы используют различные фильтры для идентификации заданных форм.

Авторы [7] приводят способы выделения признаков формы. Для их описания могут служить глобальные характеристики форм или различные методы статистического распознавания форм. При структурных моделях форма описывается как совокупность примитивных элементов и пространственных отношений между ними. В некоторых случаях форма представляется в виде графа, что позволяет сопоставлять изображения независимо от двумерных преобразований (поворот, масштабирование). Однако при осуществлении поиска в базах данных инвариантность к повороту необходима не всегда. Примером объектов, при распознавании которых заранее известна ориентация, служат деревья, автомобили, здания.

В качестве простой характеристики формы иногда применяют гистограммы формы, позволяющие выделить область интереса на обрабатываемой фотографии путем исключения частей, не соответствующих форме из запроса. Методы на основе эскизного сопоставления рассматривают форму как совокупность нескольких границ, что делает описание более полным и точным.

Далее будут описаны основные методы сегментации, применяемые в задачах поиска по форме.

1.2.1 Пороговая обработка

Пороговая обработка широко применяется для сегментации изображений, характеризующихся наличием однородных участков с различной средней яркостью. Наиболее распространенным методом является бинарная сегментация, в случае которой области изображения могут относиться к одному из двух участков. Преобразование исходного изображения производится с помощью выражения (1.10).

$$y(i_1, i_2) = \begin{cases} y_0, & \text{при } x(i_1, i_2) \leq x_0, \\ y_1, & \text{при } x(i_1, i_2) > x_0, \end{cases} \quad (1.10)$$

где $x(i_1, i_2)$ – точка исходного изображения;

$y(i_1, i_2)$ – точка выходного изображения;

y_k – уровень яркости 1-го и 2-го участков выходного изображения;

x_0 – значение порога.

Важным моментом при реализации метода является реализация автоматического определения значения порога. В источнике [8] описывается базовый метод на основе анализа гистограммы исходного изображения. Для определения оптимального порога рассчитывается дискриминантная функция $\eta(x_0^*)$ по формуле (1.8), аргумент x_0^* которой (предполагаемое значение порога) при максимуме функции считается искомым значением параметра сегментации x_0 .

$$\eta(x_0^*) = \frac{\sigma_B^2(x_0^*)}{\sigma_T^2}, \quad (1.11)$$

где σ_B^2 – межклассовая дисперсия, определяемая выражением (1.12);

σ_T^2 – дисперсия для всего изображения, определяемая выражением (1.13).

$$\sigma_B^2(x_0^*) = \frac{[\mu(255)p(x_0^*) - \mu(x_0^*)]^2}{p(x_0^*)[1 - p(x_0^*)]}, \quad (1.12)$$

где $\mu(255) = \sum_{x=0}^{255} xw(x)$ – математическое ожидание для всего изображения;

$p(x_0^*)$ – вероятность того, что произвольная точка x_0^* принадлежит области 1;

$\mu(x_0^*) = \sum_{x=0}^{x_0^*} xw(x)$ – ненормированное математическое ожидание для области

1.

$$\sigma_T^2(x_0^*) = \sum_{x=0}^{255} |x - \mu(255)|^2 w(x), \quad (1.13)$$

где $w(x)$ – значение гистограммы в точке x .

Вероятность $p(x_0^*)$ определяется по формуле (1.14).

$$p(x_0^*) = \sum_{x=0}^{x_0^*} w(x) \quad (1.14)$$

На рисунке 1.4 изображен пример гистограммы для выбора оптимального значения порога.

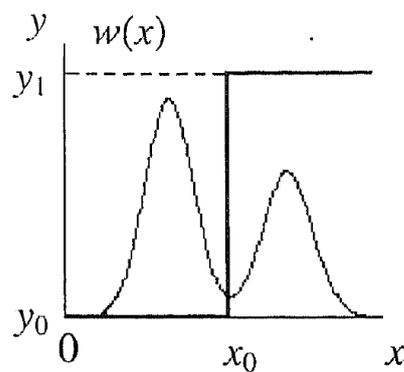


Рисунок 1.4 – Пример гистограммы при выборе порога

Результаты анализа описанного выше метода представлены на рисунке 1.5. Для сравнения на рисунке 1.6 приведены гистограмма интенсивности исходного изображения и дискриминантная функция $\eta(x_0^*)$.

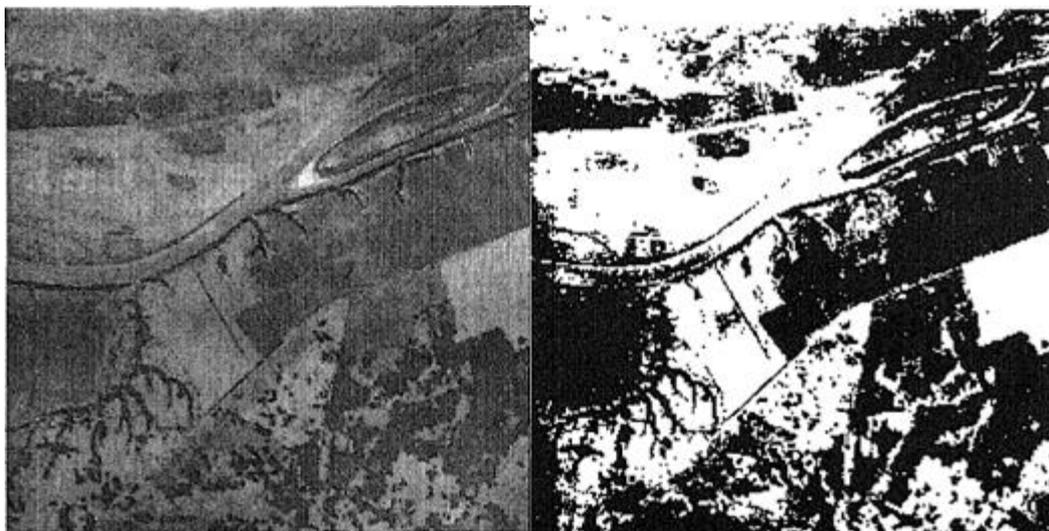


Рисунок 1.5 – Результат пороговой сегментации спутникового снимка

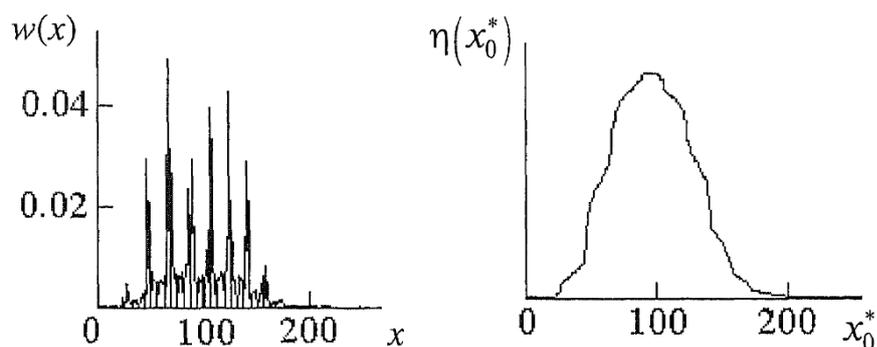


Рисунок 1.6 – Гистограмма исходного изображения (рисунок 1.5) и график дискриминантной функции

Применение алгоритмов пороговой сегментации ограничивается обработкой простых кадров, а в случае с изображениями с неровным освещением, тенями и помехами эффективность метода резко ухудшается. В [8] предлагается разбивать гистограмму на более мелкие части для определения ярко выраженных экстремумов, а также применять методы анализа взвешенных значений этих точек.

1.2.2 Марковская фильтрация

Статистический метод сегментации, основанный на Марковской фильтрации, подразумевает принятие решения в каждой точке на основе анализа изображения в ее окрестностях. В [8] рассматривается некаузальный вариант метода, когда окрестность составляют пересечение столбца и строки исследуемой точки. На рисунке 1.7 приведен пример сегментации данным методом [7].

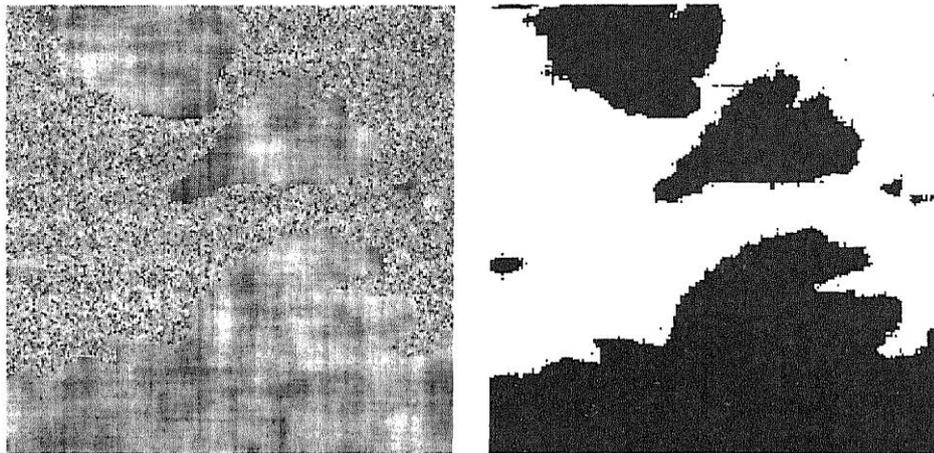


Рисунок 1.7 – Марковская сегментация текстур

В [9] алгоритм описывается в терминах Марковских случайных полей, представляющих собой графическую модель, которая задается графом $G = (V, E)$. Энергия модели записывается выражением (1.15).

$$E(X) = \sum_{i \in P} D_i(x_i) + \sum_{(i,j) \in E} V_{ij}(x_i, x_j), \quad (1.15)$$

где P – множество переменных;

E – система соседства;

V – бинарные потенциалы;

D – унарные потенциалы;

$X = (x_{ij}), x_{ij} \in \{1, \dots, K\}$ – метки.

Задача сегментации сводится к восстановлению наиболее вероятных скрытых переменных X по наблюдаемым переменным Y [<http://www.machinelearning.ru/wiki/images/6/61/MRFTutorial2.pdf>] (1.16).

$$X_{MP} = \arg \max_X P(X | Y), \quad (1.16)$$

В [10] предложены различные алгоритмы восстановления скрытой компоненты. Базовый алгоритм включает 3 прохода по графу G . Обработка начинается с корня t^* , при этом задается априорное распределение $q_{t^*}(x_{t^*}), x_{t^*} \in \Omega$. Нисходящим просмотром для всех узлов $t \in T$ по формуле (1.17) вычисляются априорные распределения.

$$q_s(x_s) = \sum_{x_t \in \Omega} q_s(x_s | x_t) q_t(x_t), x_s \in \Omega, s \in T_{(t)}^{+0}, \quad (1.17)$$

где s, t – ребра графа G .

Следующие этапы включают в себя восходящий и нисходящий просмотры для вычисления апостериорных маргинальных распределений классов по формуле (1.18), на основании которых принимаются решения о классах.

$$p_t(x_t | Y_t^+) \propto p_t(x_t | Y_{(t)}^+) p_t(x_t | y_t), x_t \in \Omega, t \in T, \quad (1.18)$$

$$p_t(x_t | Y_{(t)}^+) \propto \prod_s \sum_{x_s \in \Omega} p_t(x_t | Y_t^+) \frac{q_s(x_s | x_t)}{q_s(x_s)}, s \in T_{(t)}^{+0},$$

где Y_t^+ – поддереву с корнем в y_t , включая его.

Сегментация на основе Марковских полей особенно актуальна, когда необходимо обработать изображение, содержащее в себе текстуры, поскольку пороговая сегментация в подобном случае неприменима.

1.2.3 Граничные алгоритмы

Алгоритмы данной группы заключаются в нахождении точек на границах областей. Для распознавания границ часто используются характеристики яркости, текстуры и градиент. Само определение границы содержит понятие локальных перепадов яркости на изображении, которые могут быть линейными и пошаговыми. В случае линейной неоднородности на изображении присутствует некоторый фрагмент, характеризующийся резким скачком или падением яркости. При пошаговом разрыве параметр яркости меняет свое значение постепенно, порой доходя до диаметрально противоположного значения [8].

В общем случае можно выделить базовую последовательность этапов в алгоритмах выделения границ:

- фильтрация исходного изображения для снижения влияния шумов;
- усиление областей с перепадом яркости;

- выделение точки в границу с использованием порогового значения;
- локализация местоположения и направления.

В качестве способов выделения границ часто применяются операторы Собела, Канни, Робертса [8]. Детектор краев Кэнни является наиболее популярным и эффективным оператором [7]. После сглаживания полутонового изображения формируются протяженные контурные сегменты путем обнаружения соседних пикселей с большими значениями модуля градиента. Для сравнения эффективности метода приведем результаты выделения краев с помощью операторов Кэнни и Робертса (рисунок 1.8).

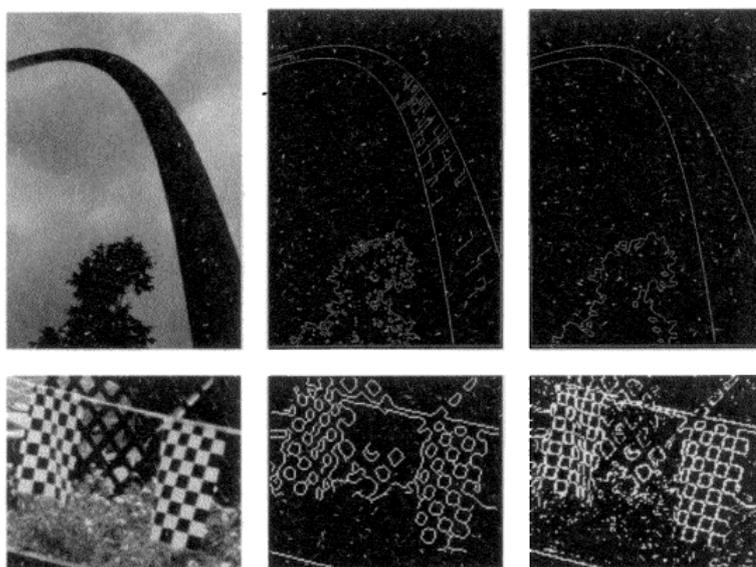


Рисунок 1.8 – Пример выделения границ на исходном изображении (слева) с помощью оператора Кэнни при пороговом значении $\sigma = 1$ и $\sigma = 4$ (посередине сверху и снизу соответственно) и оператора Робертса с параметром отбора равным 20% (справа)

1.2.4 Алгоритм разрастания регионов

Алгоритм разрастания регионов является методом автоматической сегментации, который напрямую учитывает пространственное расположение точек. В частности, он используется для выделения областей из изображения для последующего их распознавания как отдельных объектов. Однако с

помощью данного метода можно также получить разбиение на сегменты всего изображения, если применять алгоритм последовательно или одновременно для нескольких регионов.

Разрастание регионов – это подход к сегментации изображений, в котором рассматриваются соседние пиксели и добавляются в одну область, если не обнаруживается граница. Данный процесс является итеративным для каждого граничного пикселя в регионе.

Основная идея метода разрастания регионов заключается в следующем. На первом шаге выбираются центры регионов, выбор которых основан на некотором критерии пользователя. Например, пиксели, принадлежащие определенному интервалу интенсивности, или пиксели, равномерно распределенные на координатной плоскости. Начиная с выбранных центров, регионы разрастаются путем присоединения соседних точек.

Когда ни один пиксель изображения не может быть присоединен ни к одному из регионов, процесс разрастания регионов останавливается.

Точка присоединяется к региону на основании некоторого критерия, например, близость точки к центру региона, близость к соседней точке, присоединенной к региону на предыдущем шаге или же стоимость кратчайшего пути от центра региона до текущей точки и т.п.

На рисунке 1.9 представлен пример применения метода разрастания регионов к полутоновому изображению. Исходный снимок находится слева, справа – обработанное изображение с порогом 225~255.

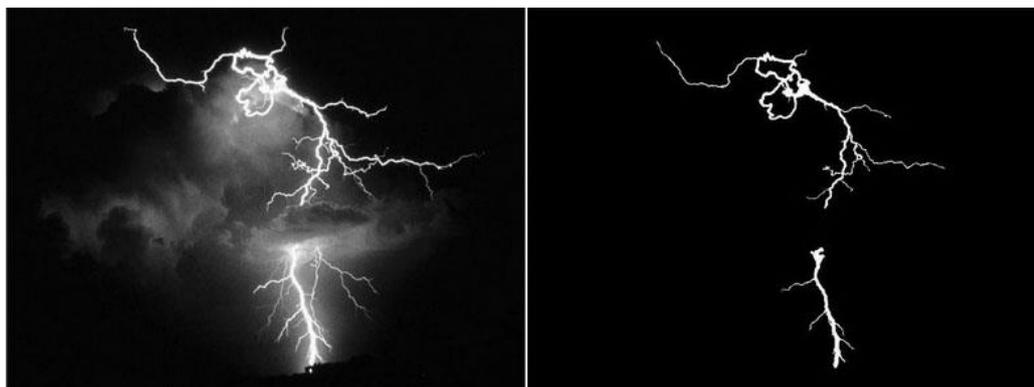


Рисунок 1.9 – Применение метода разрастания регионов

Опишем простейшую стратегию выбора центров регионов (зёрен) и разрастания регионов, которая заключается в сканировании анализируемого изображения слева направо, сверху вниз.

Рассмотрим i -й шаг алгоритма разрастания регионов. Обозначим рассматриваемую на текущей итерации точку как A , а точки, принадлежащие сформированным регионам, слева и выше от точки A , соответственно, B и C .

Обозначим яркость точки A как $I(A)$, средние яркости областей, которым принадлежат точки B и C как $I_{avg}(B)$ и $I_{avg}(C)$, а пороговое значение, которое задает чувствительность метода при создании нового сегмента как δ . Тогда условия, проверяемые на i -м шаге, выглядят следующим образом:

1. Если $|I(A) - I_{avg}(B)| > \delta \ \& \ |I(A) - I_{avg}(C)| > \delta$, то создаем новую область и присоединяем к ней пиксель A .
2. Если $|I(A) - I_{avg}(B)| \leq \delta \ \text{хор} \ |I(A) - I_{avg}(C)| \leq \delta$, то создаем новую область и присоединяем к ней пиксель A .
3. Если $|I(A) - I_{avg}(B)| \leq \delta \ \& \ |I(A) - I_{avg}(C)| \leq \delta$, то проверяем:
 1. Если $|I_{avg}(B) - I_{avg}(C)| \leq \delta$, то сливаем области B и C .
 2. Если $|I_{avg}(B) - I_{avg}(C)| > \delta$, то добавляем пиксель A к тому классу, отклонение от которого минимально [5].

В рассмотренном выше примере алгоритма оцениваются яркости точек, однако в качестве оцениваемых признаков также часто используют полутоновые текстурные или цветовые признаки.

Следует отметить, что в алгоритме разрастания регионов крайне важен выбор центров регионов, определяемый пользователем. Благодаря этому, данный метод позволяет верно выделять области с одинаковыми характеристиками, указанными пользователем. Другим достоинством метода разрастания регионов является то, что он прост для понимания. Кроме того, данный алгоритм позволяет выбрать составной критерий для разделения изображения на регионы. Основным же недостатком метода является высокие вычислительные затраты.

1.3 Текстурные характеристики

В отличие от цветовых признаков, описание текстуры является более сложным процессом, поскольку учитывает не только пространственные соотношения, но и совпадение распределения цветов (или интенсивности), при котором значения цветов двух кадров могут отличаться [7]. Далее рассмотрим основные способы описания текстурных характеристик.

1.3.1 Структурный способ представления текстур

Структурный подход (или синтаксический) развился после возникновения теории формальных языков в середине пятидесятых годов прошлого века. Он основан на том, что текстура составлена из регулярно или почти регулярно повторяющихся непримитивных элементов. Поэтому описание текстуры должно состоять из описания непримитивных элементов и правил их размещения. Далее, производя поиск соответствующих компонент на любом другом изображении, происходит распознавание уже известной текстуры.

В данном методе представления текстур выделяют две стадии: стадия обучения и стадия распознавания. Стадия обучения, в свою очередь, состоит из следующих этапов:

1. Сегментация. На этом этапе на эталонных текстурах выделяются однородные связные области. В результате имеем множество однородных связных областей изображений:

$$X = \{X_1, X_2, \dots, X_U\}. \quad (1.19)$$

В общем случае может использоваться любой алгоритм сегментации, не требующий априорных сведений об изображении.

2. Кластеризация. Множество однородных связных областей всех имеющихся эталонных изображений текстур разбивается на множество классов однородных областей:

$$A = \{A_1, A_2, \dots, A_Q\}, A_i \subseteq X, \forall A_i, A_j: A_i \cap A_j = \emptyset, i \neq j. \quad (1.20)$$

При этом в общем случае может использоваться любой алгоритм кластеризации. Каждой текстуре ставится в соответствие свое множество входящих в нее классов:

$$I_r \rightarrow \{A_1 | A_2 | \dots | A_n\}, A_i \in A, r = 1, 2, \dots, L. \quad (1.21)$$

3. Построение описания текстур. Для каждой текстуры для каждой ее однородной области выполняется обход контура и строится терминальная цепочка:

$$\alpha_i = \{b_1, b_2, \dots, b_k\}^*, \quad (1.22)$$

где * обозначает операцию замыкания. Цепочки сворачиваются к правилам, которые в дальнейшем заменяются некоторым множеством правил:

$$A_j \rightarrow B_{k1}^1 B_{k2}^2 \dots B_{km}^m, \quad (1.23)$$

В результате обучения для каждой текстуры будет построена своя грамматика, которая позволит выполнять на этапе распознавания проверку на точное соответствие входной цепочки заданной грамматике.

В режиме распознавания система состоит из следующих элементов:

1. Сегментация. На этом этапе на исходном изображении выделяются однородные связные области. Получаем множество:

$$X' = \{X'_1, X'_2, \dots, X'_p\}. \quad (1.24)$$

Должен использоваться тот же алгоритм, что и на этапе обучения, с теми же самыми параметрами.

2. Классификация. Для каждого из элементов множество однородных связных областей определяется класс, которому она соответствует:

$$X'_j \rightarrow A_i. \quad (1.25)$$

Должно использоваться множество классов, полученное на этапе обучения. Алгоритм классификации, определяющий способ построения решающих правил, в общем случае может быть любым.

3. Распознавание. Для каждой однородной области входного изображения выполняется обход контура и строится терминальная цепочка:

$$\alpha_i = \{b_1, b_2, \dots, b_k\}^*. \quad (1.26)$$

Для каждой цепочки, используя стандартные алгоритмы, определяется ее принадлежность той или иной грамматике. Если цепочка принадлежит данной грамматике, то соответствующая ей однородная область составляет текстуру, которая данной грамматикой определяется.

Из описания метода следует набор требований к алгоритму сегментации:

- области должны быть однородны относительно определенных характеристик;
- смежные области должны существенно отличаться по значениям выбранных характеристик, относительно которых они считаются однородными;
- границы каждого сегмента должны быть простыми, пространственно-точными.

Один из вариантов структурного геометрического описания текстур был предложен Тусерьяном и Джейн в 1990 г. В качестве текселов

выбираются области изображений, которые можно обнаружить с помощью простых операций обработки наподобие пороговой бинаризации. Характеристики взаимосвязи текселов определяются на основе построенной для них диаграммы Вороного.

Предположим, что имеется множество обнаруженных текселов, и каждый тексел можно представить одной характерной точкой, например, центром тяжести тексела. Обозначим множество этих точек символом S . Для любой пары точек P и Q из множества S можно построить перпендикуляр, делящий пополам отрезок PQ . Этот серединный перпендикуляр разбивает плоскость на две полуплоскости. Точки одной полуплоскости лежат ближе к точке P , а точки другой – ближе к точке Q . Обозначим через $H^Q(P)$ полуплоскость точек, более близких к точке P по отношению к серединному перпендикуляру отрезка PQ . Описанную операцию можно повторить для каждой точки Q из множества S . Многоугольником Вороного для точки P называется многоугольник, содержащий все точки, более близкие к P , чем к любой другой точке множества S [1]. Многоугольник Вороного определяется выражением:

$$V(P) = \bigcap_{Q \in S, Q \neq P} H^Q(P). \quad (1.27)$$

На рисунке 1 представлены многоугольники Вороного для множества круговых текселов. Многоугольники внутренних текселов имеют вид шестиугольников, в для граничных текселов имеют иную форму.

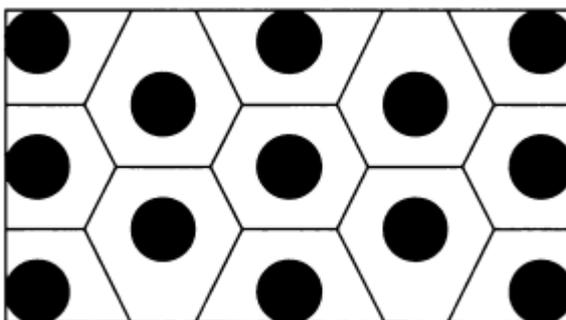


Рисунок 1.10. Диаграмма Вороного для множества круговых текселов

После того как на изображении были выделены текселы и для них определены многоугольники Вороного, можно вычислить признаки формы многоугольников и использовать эти свойства для группировки многоугольников в кластеры. Кластеры будут задавать области изображения с однородной текстурой. Шаблон в случае распространения на изображение большего размера, мог бы соответствовать некоторой области с однородной текстурой, для описания которой применимы признаки формы правильных шестиугольников.

Таким образом, в структурном подходе к описанию текстуры применяется информация о распределении таких структур изображения, как линии, края, различного рода изгибы и т.п., в отличие от статистического подхода, который использует информацию о каждом пикселе. При этом, можно сказать, что описание структур приносит некоторую выгоду, поскольку количество структур, как правило, на два или более порядка меньше, чем количество пикселей в изображении.

1.3.2 Статистический подход к описанию текстур

Как было сказано выше, основной проблемой структурного анализа текстур является сложность обнаружения текселов на естественных изображениях. Поэтому вместо обнаружения текселов по пиксельным данным полутоновых (или цветных) изображений можно вычислить количественные характеристики текстур. Статистические методы не отличаются наглядностью, однако являются более производительными, что позволяет применять их в задачах как сегментации, так и классификации [7].

Одним из методов статистического анализа является вычисление характерных текстурных признаков на основе матрицы вхождений.

Матрица вхождений – это двумерный массив S , в котором индексы строк и столбцов образуют множество V допустимых на изображении значений пикселей. Например, для полутоновых изображений V может быть

множеством допустимых значений интенсивности, а для цветных изображений – множеством возможных цветов. Значение $C(i,j)$ указывает, сколько раз значение i встречалось на изображении в некотором заданном пространственном отношении со значением j . Например, в качестве пространственного отношения может быть выбрано «значение i является правым соседом для значения j ».

Существуют две основные разновидности стандартной полутоновой матрицы вхождений. Первая называется нормированной полутоновой матрицей вхождений N_d и определяется выражением (1.28).

$$N_d[i, j] = \frac{C_d[i, j]}{\sum_i \sum_j C_d[i, j]}, \quad (1.28)$$

где $C_d[i, j]$ – это полутоновая матрица вхождения для изображения I , которая определяется в соответствие с выражением (11), в котором вектор d задает смещение между пикселем с яркостью i и пикселем с яркостью j . В выражении (1.29) через вектор d обозначен вектор перемещения (dr, dc) , где dr соответствует перемещению в направлении строк, а dc – перемещению в направлении столбцов.

$$C_d[i, j] = |\{[r, c] \mid I[r, c] = i \text{ и } I[r + dr, c + dc] = j\}|. \quad (1.29)$$

Вычисленные согласно выражению (10) значения вхождений принадлежат диапазону $[0, 1]$. Их можно рассматривать в качестве значений вероятности в матрице большего размера.

Вторая разновидность называется симметричной полутоновой матрицей вхождений S_d (1.30). В этой матрице группируются пары симметричных отношений смежности:

$$S_d[i, j] = C_d[i, j] + C_{-d}[i, j]. \quad (1.30)$$

Матрицы вхождений представляют свойства текстуры, но они неудобны для непосредственного применения при анализе изображений, например, для сравнения двух текстур. Вместо этого матрицы вхождений используются для вычисления числовых характерных признаков, которые могут служить более компактным представлением текстуры. На основе

нормированной матрицы вхождения можно вычислить следующие характерные признаки, представленные в выражениях (1.31) – (1.35):

$$\text{Энергия} = \sum_i \sum_j N_d^2[i, j], \quad (1.31)$$

$$\text{Энтропия} = - \sum_i \sum_j N_d[i, j] \log_2 N_d[i, j], \quad (1.32)$$

$$\text{Контраст} = \sum_i \sum_j (i - j)^2 N_d[i, j], \quad (1.33)$$

$$\text{Однородность} = \sum_i \sum_j \frac{N_d[i, j]}{1 + |i - j|}, \quad (1.34)$$

$$\text{Корреляция} = \frac{\sum_i \sum_j (i - \mu_i)(j - \mu_j) N_d[i, j]}{\sigma_i \sigma_j}, \quad (1.35)$$

где μ_i, μ_j – средние значения,

σ_i, σ_j – среднеквадратические ожидания сумм $N_d[i]$ и $N_d[j]$, вычисленных в направлении строк и столбцов изображения:

$$N_d[i] = \sum_j N_d[i, j], \quad (1.36)$$

$$N_d[j] = \sum_i N_d[i, j]. \quad (1.37)$$

Одна из проблем при вычислении текстурных характеристик на основе матриц смежности заключается в выборе вектора смещения d . Для выбора значения d , соответствующего наиболее выраженной структуре текстурного образа, можно использовать статистическую проверку $\chi^2[1]$. Это значение соответствует максимуму величины:

$$\chi^2(d) = \left(\sum_i \sum_j \frac{N_d^2[i, j]}{N_d[i] N_d[j]} - 1 \right). \quad (1.38)$$

Еще один подход к вычислению текстурных характеристик предполагает обнаружение различных типов текстур с помощью локальных масок. Лавс разработал энергетический подход, в котором оценивается изменение содержания текстуры в пределах окна фиксированного размера.

Для вычисления энергетических характеристик изображения используются следующие фильтры:

$$L3 \text{ (Level)} = [1 \quad 2 \quad 1 \quad]$$

$$E3 \text{ (Edge)} = [-1 \quad 0 \quad 1 \quad]$$

$$S3 \text{ (Spot)} = [-1 \quad 2 \quad -1 \quad]$$

Как видно, длина данных векторов равна трём, однако часто используются расширенные фильтры, получаемые в результате применения операции свёртки к двум из указанных выше фильтров. Таким образом, в результате свёртки образуются девять фильтров, пять из которых также используются для выделения текстурных характеристик:

$$L5 \text{ (Level)} = [1 \quad 4 \quad 6 \quad 4 \quad 1 \quad]$$

$$E5 \text{ (Edge)} = [-1 \quad -2 \quad 0 \quad 2 \quad 1 \quad]$$

$$S5 \text{ (Spot)} = [-1 \quad 0 \quad 2 \quad 0 \quad -1 \quad]$$

$$W5 \text{ (Wave)} = [-1 \quad 2 \quad 0 \quad -2 \quad 1 \quad]$$

$$R5 \text{ (Ripple)} = [1 \quad 4 \quad 6 \quad -4 \quad 1 \quad]$$

Названия векторов описывают их назначение. Вектор L5 предназначен для вычисления симметричного взвешенного локального среднего значения. Вектор E5 предназначен для обнаружения краев, S5 – для обнаружения пятен, W5 – для обнаружения волн, а R5 – для обнаружения образа в виде ряби.

В некоторых экспериментах целесообразно расширить размерность фильтров до 7 или 9. Например, фильтры длиной 7 элементов можно получить путем свертки фильтров L5-R5 с фильтрами L3-E3. Подобные фильтры будут выглядеть следующим образом:

$$Xa7 = [1 \quad 6 \quad 15 \quad 20 \quad 15 \quad 6 \quad 1 \quad]$$

$$Xb7 = [1 \quad 4 \quad 5 \quad 0 \quad -5 \quad -4 \quad -1 \quad]$$

$$Xc7 = [-1 \quad -2 \quad 1 \quad 4 \quad 1 \quad -2 \quad -1 \quad]$$

$$Xd7 = [1 \quad 0 \quad -3 \quad 0 \quad 3 \quad 0 \quad -1 \quad]$$

$$Xe7 = [1 \quad -2 \quad -1 \quad 4 \quad -1 \quad -2 \quad 1 \quad]$$

$$Xf7 = [1 \quad -4 \quad 5 \quad 0 \quad -5 \quad 4 \quad -1]$$

$$Xg7 = [-1 \quad 6 \quad 15 \quad 20 \quad -15 \quad 6 \quad -1]$$

Перечисленные выше векторы применяются для создания двумерных фильтров путём комбинации двух одномерных фильтров. Двумерные маски получаются в результате умножения пар векторов. Например, для получения маски E5L5 надо умножить вектор E5 на L5:

$$\begin{bmatrix} -1 \\ -2 \\ 0 \\ 2 \\ 1 \end{bmatrix} \cdot [1 \quad 4 \quad 6 \quad 4 \quad 1] = \begin{bmatrix} -1 & -4 & -6 & -4 & -1 \\ -2 & -8 & -12 & -8 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 2 & 8 & 12 & 8 & 2 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}$$

В результате перемножения векторов L3-S3 получается 9 двумерных масок, представленных в таблице 1.2.

Таблица 1.2 – Двумерные фильтры Лавса

<i>1</i>	<i>2</i>
Название	Описание фильтра
L3L3	Интенсивность трех соседних пикселей в горизонтальном и вертикальном направлениях.
L3E3	Распознавание краев по горизонтали и интенсивности по вертикали.
L3S3	Распознавание точек по горизонтали и интенсивности по вертикали.
E3L3	Интенсивность в горизонтальном направлении и распознавание краев по вертикали.
E3E3	Распознавание краев в обоих направлениях.
E3S3	Распознавание точек по горизонтали и краев – по вертикали.
S3L3	Интенсивность в горизонтальном направлении и определение текстуры точек в вертикальном направлении.
S3E3	Распознавание краев в горизонтальном направлении и точек – в вертикальном.
S3S3	Распознавание точек в обоих направлениях.

На первом шаге в методе Лавса устраняется влияние интенсивности освещения. Для этого по изображению перемещается небольшое окно и значения каждого пикселя вычитается локальное среднее значение. В результате данной предварительной обработки формируется изображение, на

котором средняя интенсивность каждой пиксельной окрестности близка к 0. Выбор размера окна зависит от типа анализируемого изображения.

После предварительной обработки к полученному изображению применяется каждая из шестнадцати масок 5x5. В результате формируются шестнадцать профильтрованных изображений. Обозначим через $F_k[i, j]$ результат фильтрации пикселя $[i, j]$ с использованием k -й маски. Тогда энергетическая текстурная карта E_k для фильтра k определяется выражением (1.39).

$$E_k[r,c] = \sum_{j=c-(w-1)/2}^{c+(w-1)/2} \sum_{i=r-(w-1)/2}^{r+(w-1)/2} |F_k[i,j]|, \quad (1.39)$$

где w – это размер окрестности вокруг текущего пикселя.

Каждая текстурная энергетическая карта является полноразмерным изображением, которое представляет результаты обработки входного изображения с использованием k -й маски.

После получения шестнадцати энергетических карт некоторые симметричные пары комбинируются и в результате строятся девять окончательных карт. Каждая симметричная пара карт заменяется усредненной картой. Например, маска E5L5 характеризует содержание горизонтальных краев, а L5E5 – вертикальных. Среднее двух этих карт будет характеризовать наличие на изображении краев обоих типов.

Девять изображений энергетических карт можно рассматривать в качестве одного изображения, в котором каждому пикселю соответствует вектор из 9 текстурных атрибутов. Эти текстурные характеристики можно использовать для кластеризации изображения на области с одинаковой текстурой.

На рисунке 1.11 представлены результаты сегментации изображений с помощью алгоритма кластеризации на основе областей с использованием энергетических текстурных характеристик Лавса. Слева представлены исходные изображения, справа – результат сегментации на 4 кластера.

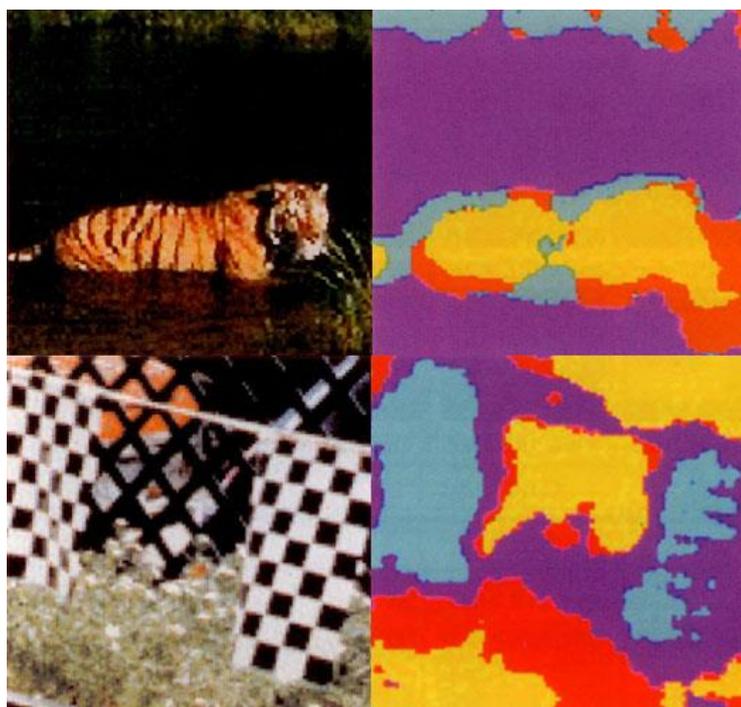


Рисунок 1.11 – Результаты сегментации с использованием энергетических текстурных характеристик Лавса

Кроме описанных выше статистических методов представления текстур, в настоящее время широко применяется подход, основанный на использовании фильтров для выделения текстурных признаков. Как известно, свертка изображения с линейным фильтром представляет изображение в другом базисе. Преимущество преобразования изображения в новый базис, которое дает его свертка с фильтром, состоит в том, что этот процесс выявляет структуру изображения. Это происходит благодаря тому, что реакция фильтра сильнее тогда, когда узор изображения на каком-либо участке похож на ядро фильтра, и слабее в противном случае.

Таким образом, можно представить текстуру изображения в виде реакции набора фильтров. Набор различных фильтров должен состоять из серий узоров – как правило, пятен и полос – при определенной последовательности масштабов.

В качестве инструмента анализа текстур часто используются фильтры Габора. Декомпозиция Габора представляет собой так называемый «направленный микроскоп» с набором ориентаций и масштабам.

Фильтры Габора принадлежат к семейству полосовых фильтров. Такие фильтры способны выявить диапазон частот сигнала в определенном промежутке и направлении. Ядра фильтра Габора выглядят как элементы базиса Фурье, умноженные на гауссианы. Таким образом, фильтры Габора дают сильную реакцию в тех точках изображения, где есть компоненты с локальными особенностями пространственной частоты и ориентации.

Преобразование Фурье импульсной передаточной характеристики фильтра Габора является сверткой преобразований Фурье гармонической функции и гауссиана [7]:

$$G(x', y') = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cdot \cos\left(2\pi \frac{x'}{\lambda} + \phi\right). \quad (1.40)$$

В формуле (1.40) параметры x' и y' определяются в соответствие с выражением (1.41) и определяют радиус размытия для фильтра Гаусса. Они представляют собой расстояние от пикселя $[x', y']$ до центрального пикселя окрестности $[x, y]$ на выходном изображении, формирования которого используется фильтр. σ – стандартное отклонение распределения Гаусса или дисперсия; γ – это коэффициент сжатия, характеризующий эллиптичность функции Габора; λ - длина волны множителя-косинуса, а ϕ – сдвиг фаз в градусах.

Вращение фильтра на θ градусов описывается выражениями (1.41) и (1.42).

$$x' = x \cdot \cos\theta + y \cdot \sin\theta, \quad (1.41)$$

$$y' = -x \cdot \sin\theta + y \cdot \cos\theta. \quad (1.42)$$

Таким образом, изменяя угол θ и размер окна фильтра, можно получить фильтры для определения полос различного масштаба и направления. При этом, как правило, θ принимает следующие значения:

$$\left\{0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}\right\}.$$

Экспериментальным путем было доказано, что коэффициент сжатия γ изменяется в пределах от 0,23 до 0,92. Полоса пропускания фильтра Габора определяется значением параметров λ и σ в соответствии с формулой (1.43):

$$b = \log_2 \frac{\frac{\sigma}{\lambda} \pi + \sqrt{\frac{\ln 2}{2}}}{\frac{\sigma}{\lambda} \pi - \sqrt{\frac{\ln 2}{2}}}, \quad (1.43)$$

$$\frac{\sigma}{\lambda} = \frac{1}{\pi} \cdot \sqrt{\frac{\ln 2}{2}} \cdot \frac{2^b + 1}{2^b - 1}. \quad (1.44)$$

Сдвиг фаз ϕ равняется 0° или 180° .

На рисунке 1.12 изображены фильтры Габора пяти различных размеров и ориентаций.

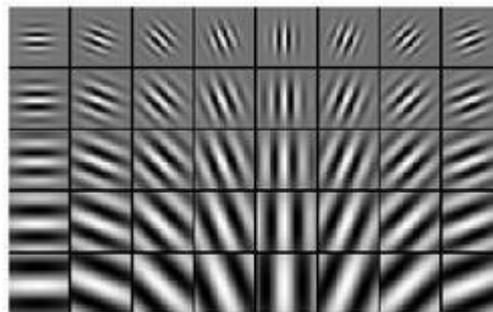


Рисунок 1.12 – Примеры фильтров Габора

Результаты применения фильтров Габора изображены на рисунке 1.13. Исходные изображения представлены слева, изображения после фильтрации – справа. При этом размер окна фильтра Габора был выбран равным 3, длина волны $\lambda = 1/16$, а ориентация фильтра $\theta = \pi/3$.

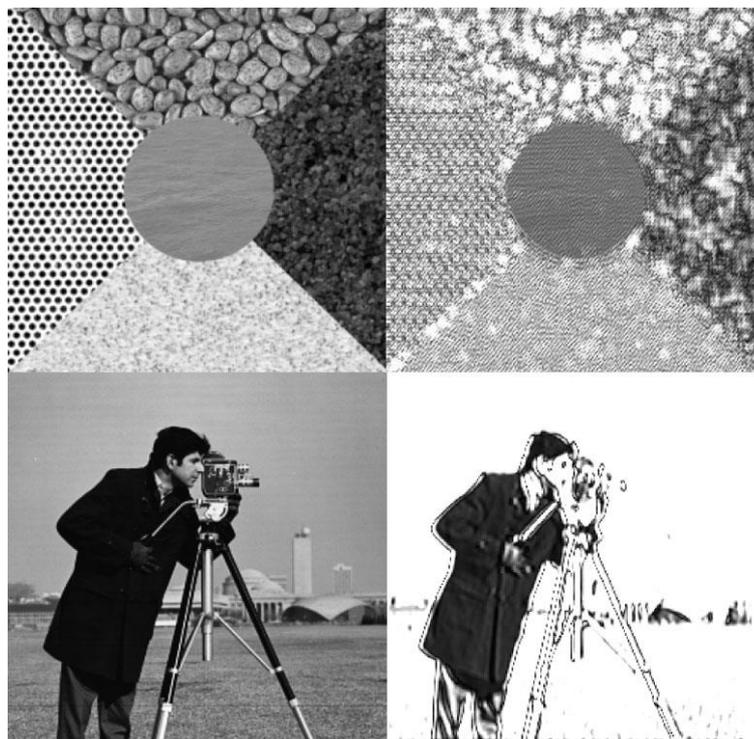


Рисунок 1.13 – Результаты применения фильтров Габора для полутоновых изображений.

Таким образом, фильтры Габора позволяют выделить текстурные характеристики для последующей сегментации. Этот процесс можно представить следующими этапами:

- генерация фильтров Габора;
- фильтрация исходного изображения с помощью полученных на предыдущем этапе фильтров;
- получение текстурных признаков;
- кластеризация каждого пикселя в пространстве текстурных характеристик.

2 Реализация метода поиска текстур в видеопотоке

В главе 1 были рассмотрены некоторые методы выделения текстурных характеристик. Для описания текстур были выбран статистический подход на основе энергетических характеристик Лавса, поскольку в случае с обработкой видеоданных важным фактором является производительность алгоритма. При расчете энергетических характеристик многие данные требуется определить один раз в самом начале работы программы без последующей необходимости их пересчета при изменении параметров алгоритма.

2.1 Выбор и обоснование среды разработки

Разработка программы для поиска текстур ведется для операционной системы Windows. В качестве среды реализации использовалась среда Visual Studio 2013, благодаря распространенности которой существует широкий спектр инструментов для работы с графической информацией, включая видеофильмы.

Наиболее популярной является библиотека с открытым исходным кодом OpenCV (Open Source Computer Vision Library), реализующая алгоритмы компьютерного зрения, специальные методы обработки изображений и методы общего назначения. Основным достоинством OpenCV является возможность бесплатного использования библиотеки, как в академических, так и коммерческих целях, а большое количество документации и обучающих материалов позволяет более эффективно решать поставленные задачи. Осуществляется поддержка большинства современных операционных систем, включая Windows. Однако используемый для разработки программы язык C# 5.0 не предусмотрен в качестве средства работы с OpenCV.

Поэтому выбор графической библиотеки стоял между тремя наиболее распространенными библиотеками-обертками для OpenCV, позволяющими использовать ее методы в программах, написанных на C#:

- Emgu CV;
- OpenCvDotNet;
- OpenCvSharp;

Для решения поставленной задачи была выбрана первая обертка, поскольку ее популярность среди программистов обеспечила наиболее полную документацию. В следующем разделе будут рассмотрены основные особенности Emgu CV.

2.2 Библиотека машинного зрения Emgu CV

Являясь кроссплатформенной оберткой популярной графической библиотеки OpenCV, Emgu CV обеспечивает доступ программистов C# практически ко всем функциям основной библиотеки. В настоящее время самой последней стабильной версией является Emgu.CV-2.4.2, но ведется активная отладка бета-версии Emgu.CV-2.4.9.

Подробная документация о продукте доступна на официальном сайте [12]. Библиотека содержит 2 уровня, первый из которых содержит аналоги базовых функций, структур и перечислений OpenCV. Второй – классы, объединяющие в себе достоинства средств работы с графикой, реализованных в фреймворке .NET.

В некоторых случаях Emgu CV позволяет решать некоторые задачи двумя подходами, используя классы первого уровня или же функцию Emgu.CV.CvInvoke, которая служит для вызова функций библиотеки OpenCV. При этом при вызове используется название функции OpenCV без каких-либо изменений.

При обработке изображений важно иметь средство для работы с матрицами. В Emgu CV задание матрицы осуществляется с помощью вызова:

```
Matrix<Depth> matrix = new Matrix<Depth>(width, height);
```

Поддерживаются типы:

- Byte, SByte;
- Single (float), Double;
- UInt16, Int16, Int32.

Работа с изображениями осуществляется с помощью класса Image, который позволяет задавать глубину и цветовое пространство изображения с помощью параметров <TColor, TDepth>:

```
Image<Gray, Byte> image = new Image<Gray, Byte>( width, height);
```

Создание изображения с использованием представленного выше кода обладает рядом преимуществ по сравнению с вызовом CvInvoke.cvCreateImage. Во-первых, в случае вызова конструктора память автоматически освобождается с помощью сборщика мусора. Во-вторых, класс Image<TColor, TDepth> реализует множество полезных методов для доступа к содержимому изображений, в том числе и преобразования к другим типам.

2.3 Сравнение энергетических карт

Энергетические характеристики Лавса позволяют представить изображение таким образом, что участки с похожими текстурами будут иметь близкие значения. Важным этапом в задачах поиска текстур является сравнение энергетических карт образца и изображения, в котором осуществляется поиск.

Поскольку в результате выделения текстурных признаков мы получаем 9 энергетических карт, необходимо разработать метод объединения этих данных в одно значение, которое можно будет использовать для дальнейшего сравнения. Предложенный подход основан на вычислении среднего квадратического отклонения.

Обозначим энергетическую карту p для образца текстуры k как E . Тогда среднее значение $\bar{E}_{p,k}$ вычисляется по формуле (2.1).

$$\bar{E}_{p,k} = \frac{\sum_{i=0}^{N-1} \sum_{j=0}^{M-1} E_{i,j}}{M \cdot N}, \quad (2.1)$$

где $E_{i,j}$ – это значение энергетической характеристики Лавса в текущем пикселе;

$p = \overline{1,9}$ – количество энергетических карт;

k – количество изображений в наборе текстур.

Следующим этапом будет нахождение среднего значения для каждой энергетической характеристики по всему набору изображений (2.2).

$$\bar{E}_p = \frac{\sum_{i=0}^k \bar{E}_{p,k}}{k} \quad (2.2)$$

Для этого сумма средних значений, рассчитанных по формуле (2.1), делится на общее количество образцов.

Вычисление среднеквадратического отклонения σ_p для каждой энергетической карты производится согласно выражению (2.3).

$$\sigma_p = \sqrt{\frac{1}{k} \sum_{i=1}^k (E_{p,i} - \bar{E}_p)^2} \quad (2.3)$$

Исходное изображение, в котором осуществляется поиск, также необходимо представить в виде энергетических карт Лавса. В результате каждый пиксель изображения будет представлен девятью характеристиками, которые необходимо поочередно сравнить со значениями, полученными на предыдущем этапе, согласно выражению (2.4).

$$\bar{E}_p - 3 \cdot \sigma_p < E_{i,j,p} < \bar{E}_p + 3 \cdot \sigma_p \quad (2.4)$$

где $i = \overline{1,N}, j = \overline{1,M}, p = \overline{1,9}$.

Выражение (3.4) представляет собой правило трех сигм. Данный закон распределения случайной величины формулируется следующим образом. Практически все значения нормально распределенной величины будут

лежать в промежутке $(\bar{x} - 3\sigma; \bar{x} + 3\sigma)$. Вероятность принадлежности величины интервалу составляет 0,9973 (рисунок 2.1).

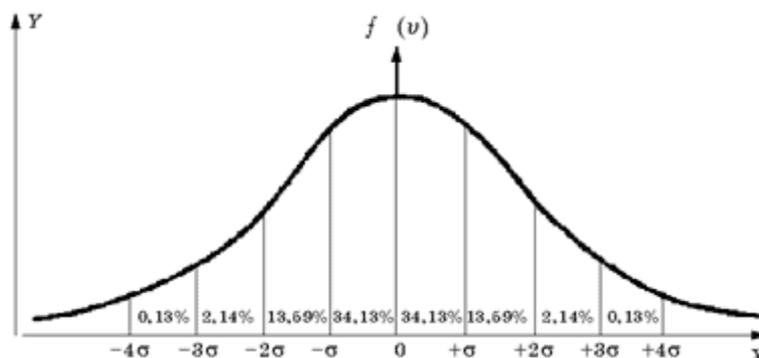


Рисунок 2.1 – График плотности вероятности нормального распределения

Отметим, что данное правило применимо в случае большое количества наблюдательных данных. В реальных задачах мы имеем среднеквадратическую ошибку ε вместо σ . Поэтому правило трёх сигм необходимо использовать с предельной осторожностью.

2.4 Комбинация текстурных и цветовых признаков

Недостатком применения метода, основанного исключительно на выделении текстурных признаков, является то, что во внимание не принимается такая важная характеристика изображений как цвет. В данном разделе описан метод комбинации данных признаков.

Цветовое пространство HSV позволяет выделить составляющую оттенка (тона), которую в дальнейшем можно использовать в качестве показателя для сравнения цветового признака образца текстуры и исходного изображения.

Обозначим значение тона пикселя i, j образца текстуры k как $H_{i,j}$. Тогда среднее значение \bar{H}_k вычисляется по формуле (2.5).

$$\bar{H}_k = \frac{\sum_{i=0}^{N-1} \sum_{j=0}^{M-1} H_{i,j}}{M \cdot N}, \quad (2.5)$$

где k – количество изображений в наборе текстур.

Если число образцов текстур больше 1, то необходимо по формуле (2.6) найти среднее значение по всем изображениям.

$$\bar{H} = \frac{\sum_{i=0}^k \bar{H}_k}{k} \quad (2.6)$$

При построении набора энергетических карт Лавса для исходного изображения, мы также определяем составляющую тона. Сравнение цветового признака осуществляется с помощью правила трёх сигм, описанного в предыдущем разделе.

Другой подход к описанию цвета заключается в построении гистограммы распределения цвета. Однако в данном случае возникает трудность реализации комбинированного метода оценки текстурных и цветовых признаков, что также может негативно сказаться на производительности алгоритма.

2.5 Реализация обработки видеоданных

Ранее рассмотренные методы применимы к статическим изображениям. Видеопоток, в общем случае, можно представить в виде последовательности таких статических кадров. Однако их последовательная обработка является очень трудоемким процессом, характеризующимся, кроме того, избыточностью. Возникает вопрос выделения ключевых кадров видеопотока, таким образом, чтобы исходный видеофайл, состоящий, например, из 1000 кадров, можно было однозначно описать с помощью 15 кадров с различными видами.

Данный процесс предобработки видеофайла состоит из двух этапов. Первый шаг заключается в разбиении исходного потока на фрагменты с похожей информацией, т.е. в поиске границ между сценами. Для реализации метода используется подход на основе сравнения гистограмм кадров в момент времени t и $t + 1$. Различие $d(t,t+1)$ между гистограммами оценивается с помощью формулы (2.7).

$$d(t,t+1) = \sum_{k=1}^2 \sum_{v=0}^V |H(I_{t+1}, c_k, v) - H(I_t, c_k, v)| , \quad (2.7)$$

где V – количество элементов в гистограмме;

c_k – компонент цветового пространства.

Формула (2.7) позволяет оценить различие между гистограммами, построенными в цветовом пространстве HSV для компонент тона и насыщенности. Если значение $d(t,t+1)$ превышает заранее заданный порог, то регистрируется новая граница между сценами.

На следующем этапе необходимо выделить ключевые кадры для каждого фрагмента видео сцены. Существует большое количество методов для решения поставленной задачи. Наиболее простым случаем является выбор кадра из середины видеофрагмента. Данный подход не отличается повышенной точностью, однако простота его реализации и отсутствие необходимости в вычислительной мощности служат обоснованием его выбора в данной работе.

Таким образом, предобработка видеофайла включает в себя 2 этапа, в результате выполнения которых объемная информация представляется некоторой последовательностью ключевых кадров. После этого к каждому ключевому кадру применяется метод поиска текстуры по образцу и в случае ее присутствия на кадре, обработке подвергается весь видеофрагмент.

3 Оценка результатов работы предлагаемого метода

В данной главе будут описаны различные параметры и модификации предложенного метода, существенным образом влияющие на его эффективность.

Поскольку построение значение энергетических карт напрямую зависит от разрешения изображения, то стоит задача определения такого разрешения изображения, при котором будут достигаться наилучший результат наравне с высокой производительностью.

Входные параметры алгоритма подбираются экспериментальным путем, поэтому в данной главе будут описаны рекомендации по их подбору.

3.1 Влияние масштаба изображения на эффективность

В качестве тестового материала для оценки качества эффективности работы программы в зависимости от качества исходного изображения использовались наборы кадров с разрешениями, представленными в таблице 3.1.

Таблица 3.1 – Разрешения тестовых кадров

Разрешение кадра, px	Поколение	Соотношение
1920×1080	Высокая четкость (Full HD)	16:9
1280×720	Высокая четкость (HD)	16:9
640×480	Стандартная четкость (SD)	4:3

В качестве входного параметра алгоритма вычисления энергетических характеристик выбран размер окрестности W равный 50 пикселям. Результаты поиска для различных размеров изображения представлены на рисунке 3.1.



Рисунок 3.1 – Результаты поиска текстуры травы на кадрах разного разрешения

Заметим, что важным фактором, влияющим на результат обработки, является соответствие между размерами образцов текстур и объектами, изображенными на кадре. Для выявления зависимости были проведены следующие тесты.

Для обработки использовалось изображение высокого разрешения, представленное на рисунке 3.2.



Рисунок 3.2 – Исходное изображение с разрешением 1280×720 px

Поиск осуществлялся по образцам различных размеров: 188×141 px; 338×190 px; 576×324 px (рисунок 4.3).



Рисунок 3.3 – Размеры образцов относительно исходного изображения

Наилучший результат был достигнут в случае соответствия разрешений исходного изображения и образца текстур (рисунок 3.4).



Рисунок 3.4 – Результаты обработки при различных размерах образцов

Уменьшить описанное выше влияние можно путем нормализации энергетических карт. Для этого во всех энергетических картах необходимо найти максимальный элемент, принимаемый за единицу, и разделить все элементы на это значение. Таким образом, энергетические характеристики будут лежать в диапазоне от 0 до 1. Недостатком такого подхода является потеря точности.

3.2 Влияние количества образцов текстур на результат

Программа предусматривает выбор числа образцов текстур. В случае количества шаблонов больше 1, вычисления среднеквадратического отклонения должны производиться в соответствии с разделом 2.3. Однако, когда пользователь руководствуется одним образцом, данный метод неприменим. Среднеквадратическое отклонение σ должно вычисляться для образца по формуле (3.1).

$$\sigma_p = \sqrt{\sum_{i=1}^N \sum_{j=1}^M (E_p[i,j] - \bar{E}_p)^2}, \quad (3.1)$$

где N, M – количество строк и столбцов в энергетической карте;

$p = \overline{1, 9}$ – количество энергетических карт;

$E_p[i,j]$ – значение энергетической характеристики в текущей точке;

\bar{E}_p – среднее значение по всей энергетической карте p , определяемое согласно формуле (3.2).

$$\bar{E}_p = \frac{\sum_{i=1}^N \sum_{j=1}^M E_p[i,j]}{N \cdot M} \quad (3.2)$$

На рисунке 3.5 представлены результаты поиска текстуры для кадра видеофайла с использованием нескольких образцов.

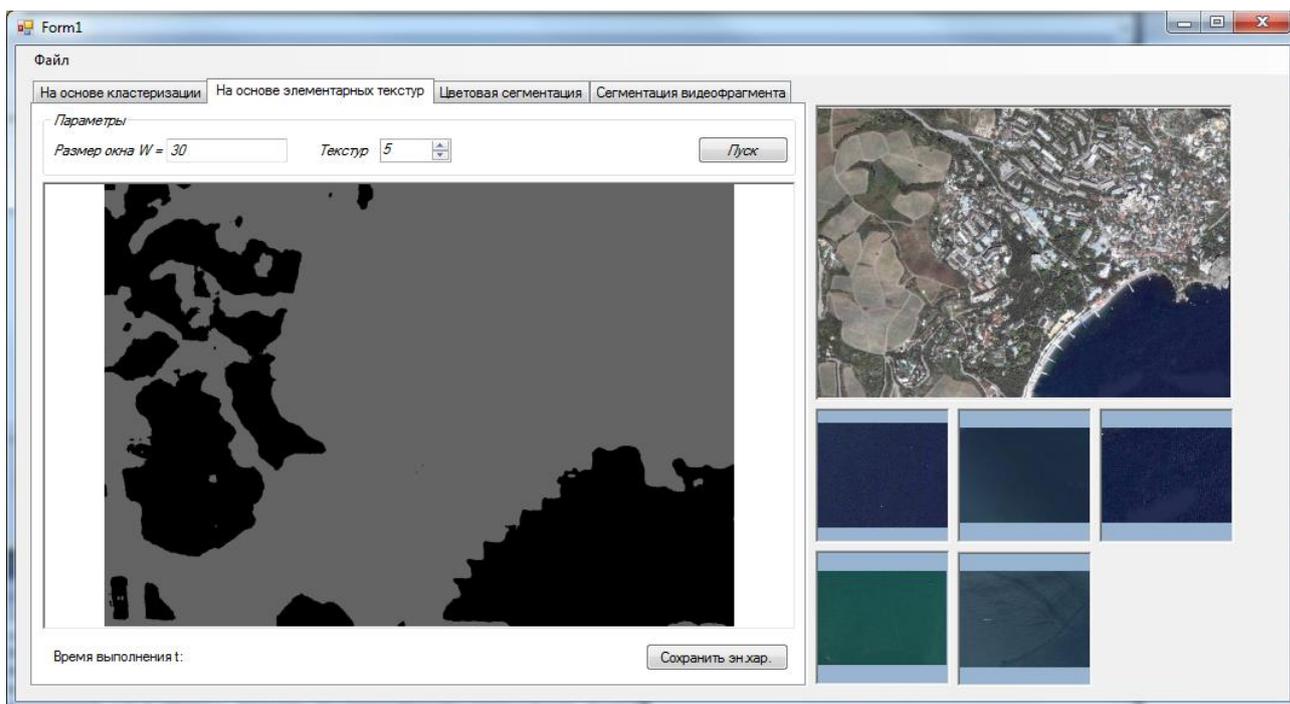


Рисунок 3.5 – Поиск текстуры на спутниковом снимке по нескольким образцам

На рисунке 3.6 приведен результат обработки того же снимка, но с использованием только одного образца поверхности воды.

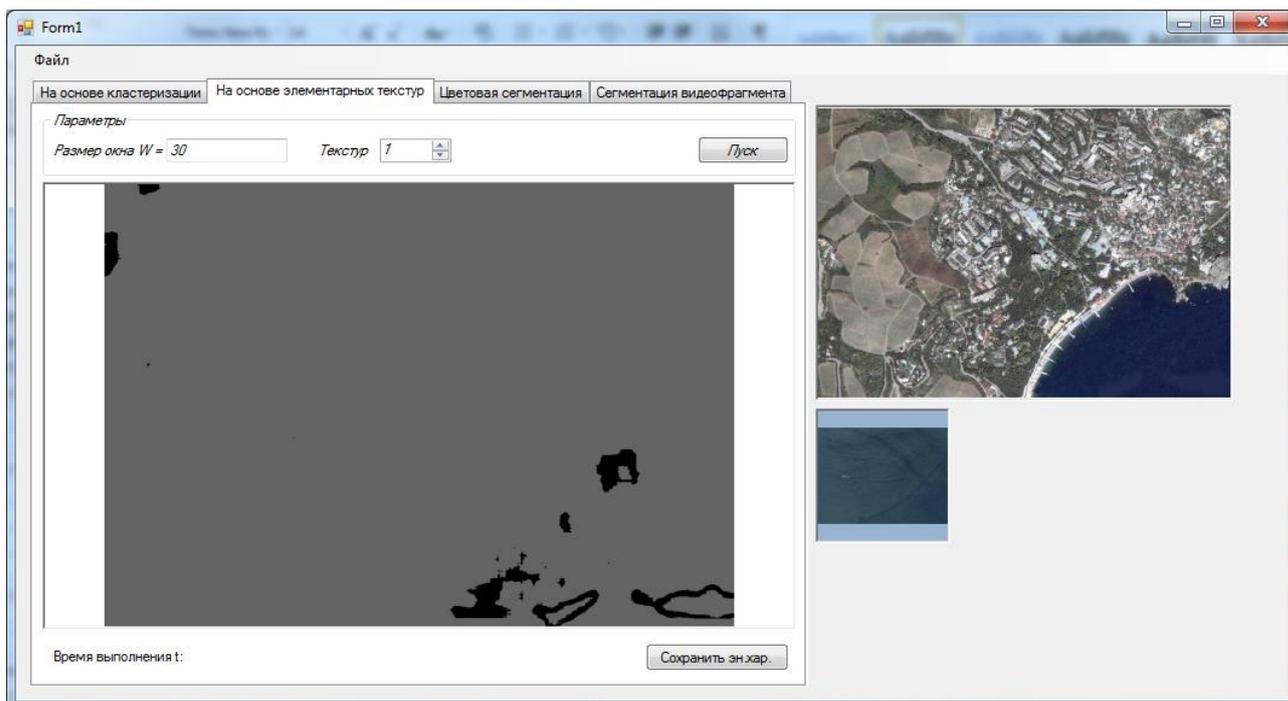


Рисунок 3.6 – Поиск текстуры на спутниковом снимке с одним образцом

Как видно, количество данных для поиска значительным образом влияет на результаты сегментации. Это объясняется тем, что для сравнения энергетических карт используется правило трех сигм, требующее наибольшее количество наблюдений. При использовании одного шаблона метод показал высокую точность поиска полностью идентичной текстуры.

3.4 Подбор оптимальных параметров

Единственным входным параметром метода, требующим подбора экспериментальным путем, является размер окрестности W . Данный аргумент определяет размеры окна вокруг центральной точки во время вычисления энергетических карт. Таким образом, учитывается не только значение текстуры в текущем пикселе, но и влияние других в области заданного размера. На рисунках 3.7 – 3.9 представлены результаты поиска текстуры на кадрах при различных значениях входного параметра W .

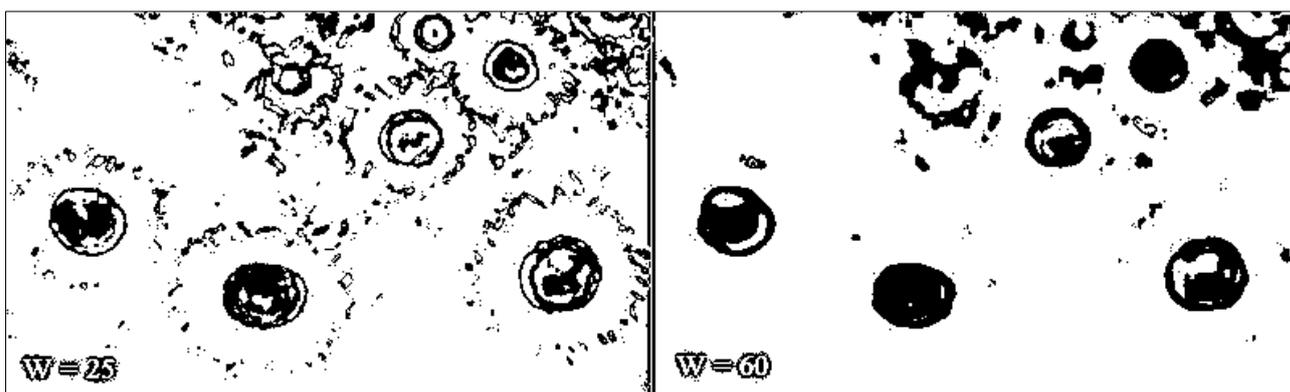


Рисунок 3.7 – Поиск на кадре разрешением 2153×1298 px

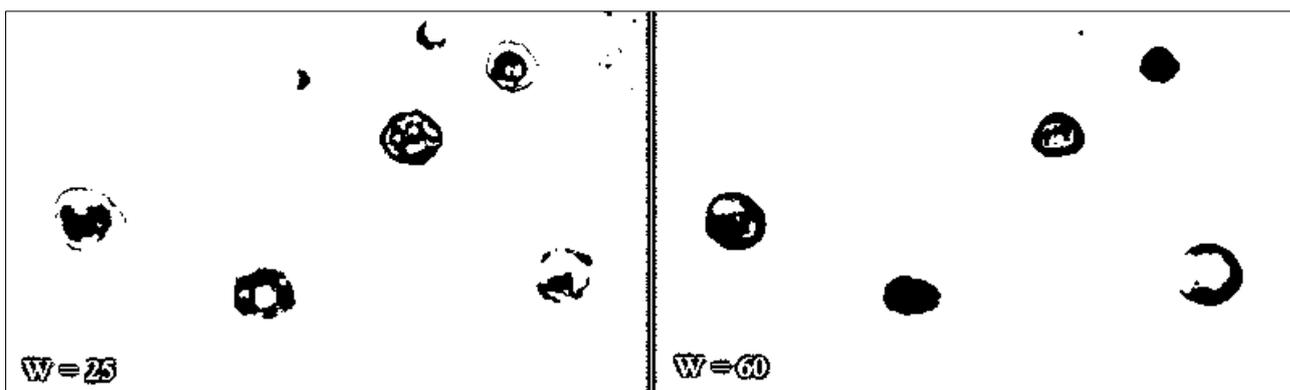


Рисунок 3.8 – Поиск на кадре разрешением 1077*649 px

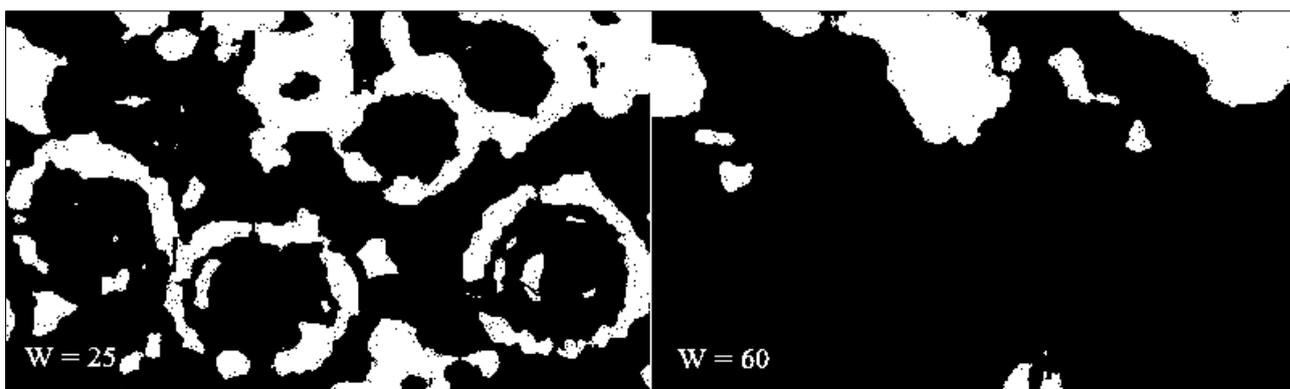


Рисунок 3.9 – Поиск на кадре разрешением 539*325 px

Таким образом, проведенные тесты показали, что существует прямая зависимость между разрешением изображения и значением параметра W . Параметр следует подбирать таким образом, чтобы размеры обрабатываемой текстуры соответствовали его величине и масштабу обрабатываемой текстуры.

**ЗАДАНИЕ ДЛЯ РАЗДЕЛА
«ФИНАНСОВЫЙ МЕНЕДЖМЕНТ,
РЕСУРСОЭФФЕКТИВНОСТЬ И РЕСУРСОСБЕРЕЖЕНИЕ»**

Студенту

Группа	ФИО
8ВМ5В	Б. С. Бокарев

Институт	Кибернетики	Кафедра	Программной инженерии
Уровень образования	Магистратура	Направление/специальность	09.04.01 Информатика и вычислительная техника

Исходные данные к разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»	
1. Стоимость ресурсов научного исследования (НИ): материально-технических, энергетических, финансовых, информационных и человеческих	
2. Нормативы расходования ресурсов	<ul style="list-style-type: none"> • Годовая норма амортизации – 37%.
3. Используемая система налогообложения	<ul style="list-style-type: none"> • Пенсионный фонд – 22%. • Фонд социального страхования – 2,9%. • Федеральный фонд обязательного медицинского страхования – 5,1%.
Перечень вопросов, подлежащих исследованию, проектированию и разработке	
1. Организация и планирование работ	<ul style="list-style-type: none"> • Приведение перечня работ по проекту. • Определение продолжительности этапов работ и составление линейного графика. • Расчет накопления готовности проекта.
2. Расчет затрат на выполнение проекта	<ul style="list-style-type: none"> • Расчет затрат на выполнение проекта по статьям расходов.
3. Оценка экономической эффективности проекта	<ul style="list-style-type: none"> • Определение типа эффекта, который производит реализация проекта и факторов экономического эффекта.
4. Оценка научно-технического уровня работ	<ul style="list-style-type: none"> • Расчет интегрального показателя научно-технического уровня работ на основе количественной оценки факторов научно-технического уровня работ.

Перечень графического материала (с точным указанием обязательных чертежей)
1. Таблица трудозатрат на выполнение проекта
2. Линейный график выполнения работ

Дата выдачи задания для раздела по линейному графику	
---	--

Задание выдал консультант

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент каф. МЕН	В. Ю. Конотопский	к. э. н.		

Задание принял к исполнению студент

Группа	ФИО	Подпись	Дата
8ВМ5В	Б. С. Бокарев		

4 Финансовый менеджмент, ресурсоэффективность и ресурсосбережение

Целью настоящего раздела магистерской диссертации является финансовая и технико-экономическая оценка показателей разрабатываемого приложения, позволяющего обрабатывать аэрокосмические снимки с беспилотного летательного аппарата. В данное обоснование включается оценка денежных затрат на исследование и разработку проекта, экономических результатов ее внедрения, а также научно-технического уровня разработки.

4.1 Организация и планирование работ

При организации процесса разработки приложения необходимо распределять уровень занятости каждого из его участников и сроки проведения отдельных этапов. Целью работы на данном этапе будет составление линейного графика проведения работ. Составим хронологическую таблицу этапов работ для исполнителя (И) и научного руководителя (НР) (таблица 4.1).

Таблица 4.1 – Перечень работ и продолжительность их выполнения

Этапы работы	Исполнители	Загрузка исполнителей
Постановка целей и задач, получение исходных данных	НР, И	НР – 60% И – 100%
Составление и утверждение ТЗ	НР, И	НР – 60% И – 100%
Подбор и изучение материалов по тематике	НР, И	НР – 10% И – 100%
Разработка календарного плана	НР, И	НР – 100% И – 10%
Разработка основных алгоритмов приложения	НР, И	НР – 30% И – 100%

Проектирование приложения	НР, И	НР – 80% И – 100%
Разработка приложения	НР, И	НР – 10% И – 100%
Оформление пояснительной записки	И	И – 100%
Оформление графического материала	И	И – 100%
Проверка работы и подведение итогов	НР, И	НР – 60% И – 100%

4.1.1 Продолжительность этапов работ

Продолжительность этапов работ рассчитывается с помощью экспертного опытно-статистического метода:

$$t_{ож} = \frac{3 \cdot t_{min} + 2 \cdot t_{max}}{5}, \quad (4.1)$$

где t_{min} – минимальная продолжительность работы, дн.;

t_{max} – максимальная продолжительность работы, дн.

Для дальнейшего построения линейного графика возникает необходимость рассчитать длительность этапов в рабочих днях, а затем перевести результат в календарные дни. Расчет продолжительности выполнения каждого этапа в рабочих днях ведется по формуле:

$$T_{рД} = \frac{t_{ож}}{K_{ВН}} \cdot K_{д}, \quad (4.2)$$

где $t_{ож}$ – продолжительность работы, дн.;

$K_{ВН}$ – коэффициент выполнения работ, учитывающий влияние внешних факторов на соблюдение предварительно определенных периодов времени, в рамках настоящей исследовательской работы установим $K_{ВН} = 1$;

$K_{д}$ – коэффициент, учитывающий дополнительное время на компенсацию непредвиденных задержек и согласование работ, примем $K_{д} = 1,05$.

Расчет продолжительности этапа проекта в календарных днях осуществляется по формуле:

$$T_{\text{КД}} = T_{\text{РД}} \cdot T_{\text{К}}, \quad (4.3)$$

где $T_{\text{РД}}$ – продолжительность выполнения этапа в рабочих днях;

$T_{\text{К}}$ – коэффициент календарности, позволяющий перейти от длительности работ в рабочих днях к их аналогам в календарных днях. Он рассчитывается по формуле:

$$T_{\text{К}} = \frac{T_{\text{КАЛ}}}{T_{\text{КАЛ}} - T_{\text{ВД}} - T_{\text{ПД}}}, \quad (4.4)$$

где $T_{\text{КАЛ}}$ – календарные дни, $T_{\text{КАЛ}} = 365$;

$T_{\text{ВД}}$ – выходные дни, $T_{\text{ВД}} = 52$;

$T_{\text{ПД}}$ – праздничные дни, $T_{\text{ПД}} = 10$.

Подставив значения в формулу 4.4, получим следующий результат:

$$T_{\text{К}} = \frac{365}{365 - 52 - 10} = 1,205$$

Таблица 04.2 – Трудозатраты на выполнение проекта

Этап	Исполнители	Продолжительность работ, дни			Трудоёмкость работ по исполнителям чел.- дн.			
		tmin	tmax	тож	ТРД		ТКД	
					НР	И	НР	И
Постановка целей и задач, получение исходных данных	НР, И	2	4	2,8	1,8	2,9	2,1	3,5
Составление и утверждение ТЗ	НР, И	2	4	2,8	1,8	2,9	2,1	3,5
Подбор и изучение материалов по тематике	НР, И	4	6	4,8	0,5	5,0	0,6	6,1
Разработка календарного плана	НР, И	2	3	2,4	2,5	0,3	3,0	0,3
Разработка основных алгоритмов приложения	НР, И	6	10	7,6	2,4	8,0	2,9	9,6
Проектирование приложения	НР, И	8	10	8,8	7,4	9,2	8,9	11,1
Разработка приложения	НР, И	30	40	34,0	3,6	5,7	4,3	3,0
Оформление пояснительной записки	И	3	5	3,8	0,0	4,0	0,0	4,8
Оформление графического материала	И	2	3	2,4	0,0	2,5	0,0	3,0
Проверка работы и подведение итогов	НР, И	2	4	2,8	1,8	2,9	2,1	3,5
Итого		61	89	72,2	1,7	3,5	6,1	8,6

Таблица 4.3 – Линейный график работ

Этап	НР	И	Февраль			Март			Апрель			Май	
			10	20	30	40	50	60	70	80	90	100	110
1	2,1	3,5	■										
2	2,1	3,5	■	■									
3	0,6	6,1		■									
4	3,0	0,3			■								
5	2,9	9,6			■								
6	8,9	11,1				■	■						
7	4,3	43,0					■	■	■	■			
8	0,0	4,8									■		
9	0,0	3,0										■	
10	2,1	3,5											■

НР – ■ ; И – ■ .

4.1.2 Расчет накопления готовности проекта

Для оценки текущего состояния проекта на основе промежуточных результатов произведем измерение накопления готовности работы. Степень готовности определяется формулой (5.5)

$$СГ_i = \frac{ТР_i^H}{ТР_{общ.}} = \frac{\sum_{k=1}^i ТР_k}{ТР_{общ.}} \quad 4.5)$$

где $ТР_i^H$ – накопленная трудоемкость i -го этапа проекта по его завершении;

$ТР_{общ.}$ – общая трудоемкость проекта;

$ТР_k$ – трудоемкость k -го этапа проекта, $k = \overline{1, i}$;

$ТР_{kj}$ – трудоемкость работ, выполняемых j -м участником на k -м этапе, здесь $j = \overline{1, m}$ – индекс исполнителя, в данной работе $m = 2$, так как в разработке проекта участвуют научный руководитель и 2 исполнителя.

Применительно к таблице 4.2 величины $ТР_{ij}$ ($ТР_{kj}$) находятся в столбцах 6 ($j = 1$) и 7 ($j = 2$). $ТР_{общ.}$ равна сумме чисел из итоговых клеток этих столбцов. Расчет $ТР_i$ (%) и $СГ_i$ (%) на основе этих данных содержится в таблице 4.4.

Таблица 4.4 – Нарастание тех. готовности работы и удельный вес этапов

Этап	$ТР_i$, %	$Г_i$, %
Постановка целей и задач, получение исходных данных	4,6	4,6
Составление и утверждение ТЗ	4,6	9,3
Подбор и изучение материалов по тематике	5,5	14,8
Разработка календарного плана	2,7	17,5

Этап	TR _i , %	Г _i , %
Разработка основных алгоритмов приложения	10,2	27,7
Проектирование приложения	16,4	44,2
Разработка приложения	38,8	82,9
Оформление пояснительной записки	3,9	92,9
Оформление графического материала	2,5	95,4
Проверка работы и подведение итогов	4,6	100

4.2 Расчет сметы затрат на выполнение проекта

Разработка информационной системы велась с использованием на домашнем компьютере и без аренды помещения. Расчет сметной стоимости выполнения проекта производится по следующим статьям затрат:

- материалы и покупные изделия;
- заработная плата;
- отчисления в Пенсионный фонд, Фонд социального страхования и Федеральный фонд обязательного медицинского страхования;
- расходы на электроэнергию (без освещения);
- амортизационные начисления;
- оплата услуг связи;
- накладные расходы.

4.2.1 Расчет затрат на материалы

К данной статье расходов отнесем стоимость материалов, которые были использованы во время разработки приложения.

Таблица 4.5 – Расчет затрат на материалы

Наименование материалов	Цена за ед., руб.	Кол-во	Сумма, руб.
Бумага для принтера формата А4	270	1 уп.	270
Итого			270

Транспортно-заготовительные расходы (ТЗР) составляют 5 % от отпускной цены материалов, тогда расходы на материалы с учетом ТЗР равны

$$C_{\text{мат}} = 270 * 1,05 = 283,5 \text{ руб.}$$

4.2.2 Расчет заработной платы

Заработная плата включает в себя сумму заработной платы исполнителей и научного руководителя, а также их премии. Данный параметр рассчитывается исходя из трудоемкости этапов и уровня занятости на каждом из них. Величина месячного оклада научного руководителя ($MO_{\text{НР}}$) получена из открытых данных, размещенных на официальном сайте Томского политехнического университета. Величина месячного оклада исполнителя ($MO_{\text{И}}$) составляет 30 тысяч рублей. Срок реализации проекта составил 3 месяца. Среднедневная заработная плата рассчитывается по формуле:

$$ЗП_{\text{дн-т}} = MO/N, \tag{4.6}$$

где MO – месячный оклад, руб.;

N – количество рабочих дней в месяц, при шестидневной рабочей неделе
 $- N = 22$

Среднедневная тарифная заработная плата научного руководителя равна

$$ЗП_{\text{дн-т}} = \frac{38800}{22} = 1763,63 \frac{\text{руб.}}{\text{раб. день}}$$

А среднедневная заработная плата разработчика равна

$$ЗП_{\text{дн-т}} = \frac{30000}{22} = 1363,63 \frac{\text{руб.}}{\text{раб. день}}$$

Затраты времени по научному руководителю и исполнителю в рабочих днях взяты из таблицы 2. Для перехода от тарифной суммы заработка исполнителя, связанной с участием в проекте, к соответствующему полному заработку необходимо будет тарифную сумму заработка исполнителя, связанной с участием в проекте умножить на интегральный коэффициент рассчитываемый по формуле:

$$K_{\text{и}} = K_{\text{пр}} \cdot K_{\text{доп.ЗП}} \cdot K_{\text{р}}, \quad (4.7)$$

где $K_{\text{пр}}$ – коэффициент премий, $K_{\text{пр}} = 1,1$;

$K_{\text{доп.ЗП}}$ – коэффициент дополнительной зарплаты, при шестидневной рабочей неделе $K_{\text{доп.ЗП}} = 1,188$, а при пятидневной рабочей неделе $K_{\text{доп.ЗП}} = 1,113$ (для НР и И);

$K_{\text{р}}$ – коэффициент районной надбавки для НР, $K_{\text{р}} = 1,3$.

Результаты вычислений представлены в таблице 4.6.

Таблица 4.6 – Затраты на заработную плату

Исполнитель	Оклад, руб./мес.	Среднедневная ставка, руб./раб. день	Затраты времени, раб. дни	Коэффициент	Фонд з/п, руб.
НР	38 800	1763,63	25	1,69	74513
И	30000	1363,63	85	1,592	184526
Итого					259039

4.2.3 Расчет затрат на социальный налог

Данная статья затрат включает в себя:

- отчисления в Пенсионный фонд России в размере 22%;
- отчисления в Фонд социального страхования в размере 2,9%;
- отчисления в Федеральный фонд обязательного медицинского страхования в размере 5,1%

Таким образом затраты на социальный налог начисляются в размере 30%:

$$C_{\text{соц}} = C_{\text{ЗП}} \cdot 0,30 \quad (4.8)$$

где $C_{\text{ЗП}}$ – размер заработной платы.

Подставив необходимые значения в формулу 4.8 получим:

$$C_{\text{соц}} = 259039 \cdot 0,3 = 77711,7 \text{ руб.}$$

4.2.4 Расчет затрат на электроэнергию

Затраты на электроэнергию рассчитываются по формуле:

$$C_{\text{эл.об.}} = P_{\text{об}} \cdot t_{\text{об}} \cdot C_{\text{э}}, \quad (4.9)$$

где $P_{\text{об}}$ – мощность, потребляемая оборудованием, кВт;

$t_{\text{об}}$ – время работы оборудования, час;

$C_{\text{э}}$ – тариф на 1 кВт·час. В данном случае, $C_{\text{э}} = 6 \text{ руб./кВт} \cdot \text{час}$.

Время работы оборудования вычисляется на основе итоговых данных таблицы 2 ($T_{\text{рд}}$) из расчета, что продолжительность рабочего дня равна 8 часов.

$$t_{\text{об}} = T_{\text{рд}} \cdot K_t, \quad (4.10)$$

где K_t – коэффициент использования оборудования по времени, $K_t = 0,9$.

Мощность, потребляемая оборудованием, определяется по формуле:

$$P_{об} = P_{ном} \cdot K_C, \quad (4.11)$$

где K_C – коэффициент загрузки. Для технологического оборудования малой мощности $K_C = 1$.

Таблица 4.7 – Затраты на электроэнергию технологическую

Наименование оборудования	Время работы оборудования $t_{об}$, час	Потребляемая мощность $P_{об}$, кВт	Затраты $\Delta_{об}$, руб.
Персональный компьютер Научного руководителя	175,6	0,1	105
Персональный компьютер Исполнителя	612,7	0,11	404,3
Итого:			508,3

4.2.5 Расчет амортизационных расходов

Расчет амортизационных расходов производится по формуле:

$$C_{ам} = \frac{N_A \cdot C_{об} \cdot t_{рф} \cdot n}{F_d}, \quad (4.12)$$

где N_A – годовая норма амортизации единицы оборудования;

$C_{об}$ – балансовая стоимость единицы оборудования с учетом ТЗР, у исполнителя – 32227 руб., научного руководителя – 29500 руб.;

$t_{рф}$ – фактическое время работы оборудования в ходе выполнения проекта, $t_{рф} = 85,1 \cdot 8 = 680,8$ часа (И), $t_{рф} = 24,4 \cdot 8 = 195,2$ часа (НР);

n – число задействованных однотипных единиц оборудования;

F_D – действительный годовой фонд времени работы соответствующего оборудования, $F_D = 298 * 8 = 2384$ часа.

N_A определяется по формуле:

$$N_A = \frac{1}{CA} \quad 4.13)$$

где CA – срок амортизации, который можно получить из постановления правительства Российской Федерации «О классификации основных средств, включенных в амортизационные группы» Для электронно-вычислительной техники CA свыше 2 лет до 3 лет включительно. В данной работе примем $CA=2,5$ года. Тогда

$$N_A = \frac{1}{2,5} = 0,4.$$

Таким образом,

$$C_{AM}(HP) = \frac{0,4 \cdot 29500 \cdot 195,2 \cdot 1}{2384} = 966,1 \text{ руб.}$$

$$C_{AM}(И) = \frac{0,4 \cdot 32227 \cdot 680,8 \cdot 1}{2384} = 3681,2 \text{ руб.}$$

Итого начислено амортизации 4647,3 руб.

4.2.6 Расчет расходов на услуги связи

Расходы на услуги связи определены наличием подключения к интернету на компьютерах, использованных в данной работе.

Ежемесячная оплата на одного разработчика составляет 400 рублей в месяц. В соответствии с таблицей 2, трудоемкость выполняемой задачи составляет три календарных месяца. Таким образом, сумма расходов на услуги связи составляет $3*400 = 1200$ руб. Расчет услуг связи для научного руководителя лежит за рамками работ по настоящему проекту, т. к. оплачивается Томским политехническим университетом.

4.2.7 Расчет прочих расходов

Прочие расходы следует принять равными 10% от суммы всех предыдущих расходов. Они находятся по формуле:

$$C_{\text{проч}} = (C_{\text{мат}} + C_{\text{ЗП}} + C_{\text{соц}} + C_{\text{эл.об.}} + C_{\text{АМ}} + C_{\text{св}}) \cdot 0,1, \quad 4.14)$$

где $C_{\text{мат}}$ – расходы на материалы, руб.;

$C_{\text{ЗП}}$ – основная заработная плата, руб.;

$C_{\text{соц}}$ – расходы на социальный налог, руб.;

$C_{\text{эл.об.}}$ – расходы на электроэнергию, руб.;

$C_{\text{АМ}}$ – амортизационные расходы, руб.;

$C_{\text{св}}$ – расходы на услуги связи, руб.

Подставив полученные выше результаты, получим:

$$C_{\text{проч}} = (283,5 + 228575 + 68572,5 + 508,3 + 4647,3 + 1200) \cdot 0,1 = 30378,7 \text{ руб.}$$

4.2.8 Расчет общей себестоимости разработки

Проведя расчет по всем статьям сметы затрат на разработку, можно определить общую себестоимость проекта «Программное обеспечение анализа изменения ландшафтного покрова по данным камеры высокого разрешения, управляемой с БПЛА с фиксированным крылом» (см. таблица 4.8).

Таблица 4.8 – Смета затрат на разработку проекта

Статья затрат	Условное обозначение	Сумма, руб.
Материалы и покупные изделия	$C_{\text{мат}}$	283,5
Основная заработная плата	$C_{\text{ЗП}}$	228575
Отчисления в социальные фонды	$C_{\text{соц}}$	68572,5
Расходы на электроэнергию	$C_{\text{эл.об.}}$	508,3
Амортизационные отчисления	$C_{\text{АМ}}$	4647,3
Расходы на услуги связи	$C_{\text{св}}$	1200
Прочие расходы	$C_{\text{проч}}$	30378,7

Итого:		334165,3
---------------	--	-----------------

Таким образом, затраты на разработку составили $C = 334165,3$ руб.

4.3 Оценка научно-технического уровня НИР

Научно-технический уровень характеризует влияние проекта на уровень и динамику обеспечения научно-технического прогресса в данной области. Для оценки научной ценности, технической значимости и эффективности, планируемых и выполняемых НИР, используется метод балльных оценок. Балльная оценка заключается в том, что каждому фактору по принятой шкале присваивается определенное количество баллов. Обобщенную оценку проводят по сумме баллов по всем показателям. На ее основе делается вывод о целесообразности НИР.

Сущность метода заключается в том, что на основе оценок признаков работы определяется интегральный показатель (индекс) ее научно-технического уровня по формуле:

$$I_{НТУ} = \sum_{i=1}^3 R_i \cdot n_i, \quad 4.13)$$

где $I_{НТУ}$ – интегральный индекс научно-технического уровня;

R_i – весовой коэффициент i -го признака научно-технического эффекта;

n_i – количественная оценка i -го признака научно-технического эффекта, в баллах.

Частные оценки уровня n_i и их краткое обоснование даны в таблице 4.9.

Таблица 4.9 Оценки научно-технического уровня НИР

Значимость	Фактор НТУ	Уровень фактора	Выбранный балл	Обоснование выбранного балла
0,4	Уровень новизны	Относительно новая	3	Систематизируются, обобщаются имеющиеся сведения, новые связи между известными факторами

Значимость	Фактор НТУ	Уровень фактора	Выбранный балл	Обоснование выбранного балла
0,1	Теоретический уровень	Разработка приложения	6	Разработка приложения для быстрого анализа лесных покровов
0,5	Возможность реализации	В течение первых лет	10	Большая прибыль при относительно малой себестоимости

Интегральный показатель научно-технического уровня для данного проекта составляет:

$$I_{НТУ} = 0,4 \cdot 3 + 0,1 \cdot 6 + 0,5 \cdot 10 = 6,8$$

По полученным данным можно сделать вывод, что проект имеет близкий к высокому средний уровень научно-технического эффекта.

4.4 Экономический эффект НИР

Количественно оценить экономический эффект от работы не представляется возможным, так как для этого требуется проводить отдельное исследование. В глобальном плане экономический эффект достигается за счет предотвращения незаконных вырубок и продажи древесных материалов, что наносит ущерб экономике региона. Для социальной сферы разработка имеет эффект в плане предотвращения загрязнений окружающей среды и сохранения естественной природы.

**ЗАДАНИЕ ДЛЯ РАЗДЕЛА
«СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»**

Студенту:

Группа	ФИО
8ВМ5В	Бокарев Борис Сергеевич

Институт	Кибернетики	Кафедра	Программной инженерии
Уровень образования	Магистратура	Направление/ специальность	09.04.01 Информатика и вычислительная техника

Исходные данные к разделу «Социальная ответственность»:

<p>1. Характеристика объекта исследования (вещество, материал, прибор, алгоритм, методика, рабочая зона) и области его применения</p>	<p>Объектом исследования являются алгоритмы, позволяющие проводить текстурный анализ спутниковых снимков и снимков с беспилотных летательных аппаратов. Результат работы – программное обеспечение, позволяющее с высокой скоростью анализировать и сравнивать спутниковые снимки на предмет поиска отклонений от заданного образца.</p>
---	--

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

<p>1. Профессиональная социальная безопасность</p> <p>1.1 Анализ вредных и опасных факторов, которые может создать объект исследования.</p> <p>1.2 Анализ вредных и опасных факторов, которые могут возникнуть на производстве при внедрении объекта исследований.</p> <p>1.3 Обоснование мероприятий по защите исследователя от действия опасных и вредных факторов.</p>	<p>1.1 Вредные производственные факторы, создаваемые объектом исследования: приложение не создает явно выраженных вредных производственных факторов</p> <p>1.2 Вредные производственные факторы, возникающие на рабочем месте:</p> <ul style="list-style-type: none"> - Неподходящий микроклимат; - Монотонность работы. - Неподходящая освещенность; <p style="text-align: center;">Опасные производственные факторы, возникающие на рабочем месте:</p> <ul style="list-style-type: none"> - Возникновение пожара. - Поражение электрическим током
<p>2. Экологическая безопасность</p> <p>2.1. Анализ влияния объекта исследования на окружающую среду.</p>	<p>2.1 Влияние объекта исследования на окружающую среду: приложение не оказывает влияние на окружающую среду, однако разрабатывается на персональном</p>

<p>2.2. Анализ влияния процесса эксплуатации объекта на окружающую среду.</p> <p>2.3. Обоснование мероприятий по защите окружающей среды.</p>	<p>компьютере, влияющем на нее.</p> <p>2.2 Влияние процесса исследования на окружающую среду:</p> <ul style="list-style-type: none"> - Утилизация неисправной электроники. - Утилизация люминесцентных ламп; - Утилизация макулатуры.
<p>3. Безопасность в чрезвычайных ситуациях</p> <p>3.1 Анализ вероятных ЧС, которые может инициировать объект исследований.</p> <p>3.2 Анализ вероятных ЧС, которые могут возникнуть на рабочем месте при проведении исследований.</p> <p>3.3 Обоснование мероприятий по предотвращению ЧС и разработка порядка действия в случае возникновения ЧС.</p>	<p>3.1 Вероятные ЧС, инициируемые объектом исследования: приложение не создает чрезвычайных ситуаций, однако разрабатывается на персональном компьютере, влияющем на нее.</p> <p>3.2 Вероятные ЧС, возникающие на рабочем месте:</p> <ul style="list-style-type: none"> - Пожар.
<p>4. Правовые и организационные вопросы обеспечения безопасности:</p> <p>4.1 Специальные (характерные для проектируемой рабочей зоны) правовые нормы трудового законодательства.</p> <p>4.2 Организационные мероприятия при компоновке рабочей зоны.</p>	<p>4.1 Описание правовых норм для работ, связанных с работой на персональном компьютере согласно следующим документам:</p> <ul style="list-style-type: none"> • Трудовой кодекс Российской Федерации" от 30.12.2001 N 197-ФЗ (ред. от 30.12.2015). <p>4.2 Влияние реализации проекта на потребителя видео-контента, как пользователя системы.</p>

Дата выдачи задания для раздела по линейному графику	
--	--

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Ассистент	Акулов Петр Анатольевич			

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8ВМ5В	Бокарев Борис Сергеевич		

5. Социальная ответственность

5.1 Введение

Разработанный программный продукт служит для обработки цифровой информации, такой как изображения и видеоданные. Основными средствами взаимодействия с программой являются элементы пользовательского интерфейса, а также периферийные устройства, подключенные к ПЭВМ. Обеспечение безопасности труда пользователей ПЭВМ является на сегодняшний день крайне актуальным вопросом, поскольку использование компьютеров становится незаменимым во многих сферах трудовой деятельности человека.

При реализации поставленной цели необходим комплексный учет всей совокупности производственных и эксплуатационных факторов на всем протяжении жизненного цикла разработки. В настоящем разделе представлено описание вредных и опасных факторов, создаваемых объектом исследования в процессе разработки, а также мероприятий, обеспечивающих производственную безопасность.

При проведении теоретических изысканий, анализа и структурирования полученных данных, проектирования и программирования приложения использовались средства вычислительной техники: персональный компьютер, серверное оборудование. Данный фактор накладывает ограничения, вызванные вредностью для человека процесса разработки. В результате у него могут начаться проблемы со здоровьем, что приведет к снижению производительности труда.

5.2 Описание рабочего места

При создании, описании алгоритмов и разработке программного обеспечения вся исследовательская деятельность проводилась в помещении Института кибернетики Томского политехнического университета. Рабочее помещение оснащено персональными компьютерами (ПК), компьютерными столами, компьютерными креслами, проекционным экраном и проектором, кондиционером, а также системой противопожарной безопасности: огнетушителями и противопожарной сигнализацией, оснащенной датчиками дыма.

В помещении нет источников повышенной, либо особо повышенной опасности. Оно относится к классу помещений без повышенной опасности. Схема помещения, в котором происходили проектирование и разработка программного продукта представлена на рисунке 5.1.

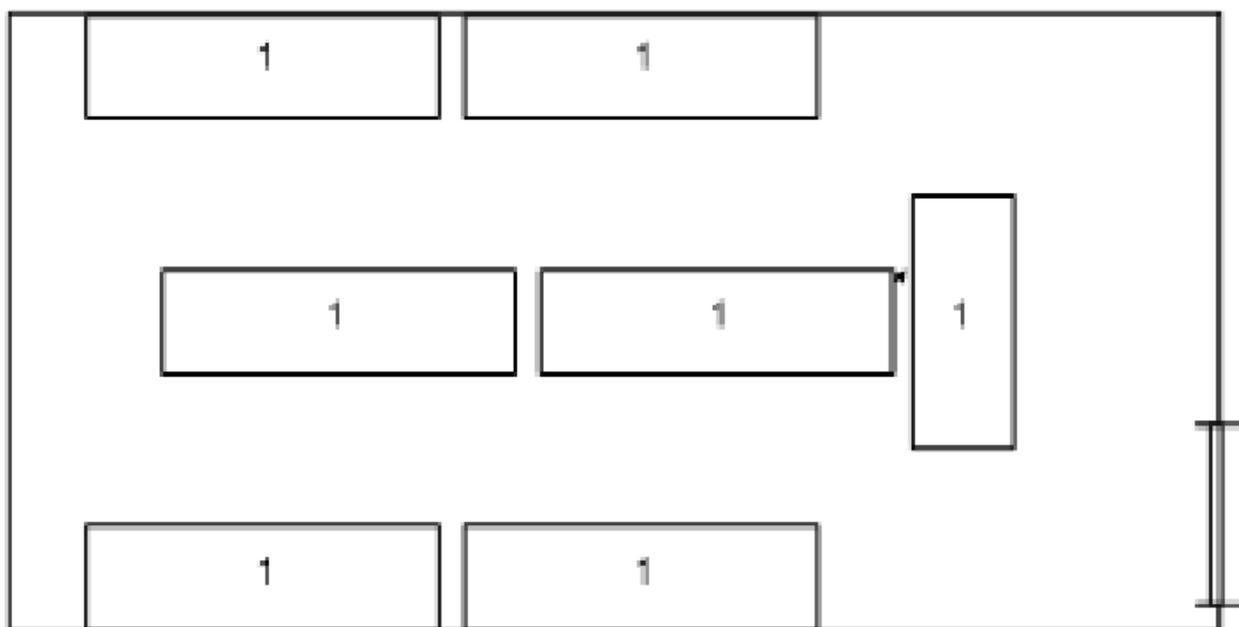


Рисунок 5.1 – Схема помещения

Рабочие пространства отдельных разработчиков расположены согласно «Гигиеническим требованиям к персональным электронно-вычислительным машинам и организации работы» (СанПиН 2.2.2/2.4.1340–03). Расстояние между рабочими столами, оснащенными видеомониторами составляет 2 м, а расстояние между боковыми поверхностями видеомониторов - около 1,2 м.

Согласно той же нормы каждое рабочее место пользователей ПК с монитором на базе плоских дискретных экранов (жидкокристаллические, плазменные) должно составлять 4,5 м² [1]. Однако в данном случае имеющиеся требования не выдерживаются в полном объеме: расстояние между рабочими столами в среднем составляет порядка 1 м. При этом расстояние между боковыми поверхностями видеомониторов соблюдено полностью.

Для соблюдения всех требований рассматриваемой нормы рекомендуется снизить количество рабочих мест в помещении.

Проведем анализ влияния вредных и опасных факторов, а также возможности возникновения чрезвычайных ситуаций. Конечная цель анализа – безопасность сотрудников и имущества Института кибернетики.

5.3 Производственная и эксплуатационная безопасность

5.3.1 Электробезопасность

В процессе разработки сотрудники должны соблюдать нормы электробезопасности. Основной риск заключается в контакте человека с частями электроприборов, находящимися под напряжением в процессе эксплуатации или профилактики. Стоит учитывать риски замыкания электрической цепи через тело человека в следствие повышенного значения напряжения, а также повышенный уровень статического электричества, повышенную напряженность магнитного и электрического полей.

Все помещения, в зависимости от условий, делятся на помещения:

- особо опасные;
- с повышенной опасностью поражения электрическим током;
- без повышенной опасности поражения электрическим током.

Возможные последствия для тела человека поражения электрическим током приведены в таблице 5.1.

Таблица 5.1 – Возможные воздействия на тело человека электрического тока

Вид воздействия	Вид электротравмы	Следствие
Электролитическое	Электрический удар	Разложение крови и других жидкостей, нарушение их состава
Термическое	Электрический ожог	Ожоги отдельных участков тела, нагрев внутренних органов
Биологическое	Механические повреждения	Судорожное сокращение мышц, Разложение и возбуждение живых тканей

Настоящий проект разрабатывался в помещении без повышенной опасности (класс 01 по ГОСТ Р 12.1.019-2009 [2]). Данный вид помещений является сухим (влажность 40–50%), средняя температура составляет 18-20° С. Электрическая проводка выполнена с использованием проводов без усиленной изоляции, пол является токонепроводящим.

Выполнение исследовательского проекта проходило в помещении, в котором используются приборы, потребляющие напряжение 220 В переменного тока с частотой 50 Гц. Данные характеристики напряжения являются опасными для жизни.

Рекомендуемые мероприятия, направленные на улучшение электробезопасности:

- не нарушать целостность источников питания;
- отключение электрооборудования в процессе его ремонта или профилактики;
- во время грозы необходимо убедиться в наличии молниеотвода, а также факте заземления всех розеток в помещении.

5.3.2 Микроклимат

Микроклимат является одним из важных условий труда. Его основные параметры регулируются СанПиН 2.2.4.548-96 [3] и ГОСТ 12.1.005-88 [4]. Персональные компьютеры и серверное оборудование выделяют большое количество тепловой энергии. В купе с недостатком воздуха в рабочей зоне данный фактор оказывает негативное влияние на человека. Психоэмоциональная обстановка в коллективе, работающем в помещении может потерять устойчивость и вызывать конфликты. Также велик риск роста болезнетворных бактерий, вызывающих различные заболевания. При недостаточной влажности воздуха слизистые оболочки человека постепенно осушаются. Данный фактор может приводить к повышенному накапливанию вредных бактерий и вирусов в дыхательных путях, затруднять работу головного мозга и других органов.

С другой стороны, повышенная влажность способствует активному размножению бактерий, плесени и грибков, что в свою очередь может привести к возникновению аллергических и хронических заболеваний дыхательных путей человека.

В санитарных нормах установлены оптимальные (таблица 5.2) и допустимые (таблица 5.3) величины параметров микроклимата, создающие комфортные условия.

Таблица 5.2 – Оптимальные величины параметров микроклимата для оператора ПЭВМ [3]

Пе риод года	Темпера тура воздуха, °С	Температ ура поверхностей, °С	Относит ельная влажность, %	Скорость движения воздуха, м/с
Холодный	22–24	21–25	60–40	0,1
Теплый	23–25	22–26	60–40	0,1

Таблица 5.3 – Допустимые величины параметров микроклимата для оператора ПЭВМ [3]

Период года	Температура воздуха, °С		Температура поверхностей, °С	Относительная влажность, % воздуха,	Скорость движения воздуха, м/с, для диапазона температур воздуха	
	ниже оптимальных величин	выше оптимальных величин			ниже оптимальных величин, не более	выше оптимальных величин, не более
Холодный	20,0–21,9	24,1–25,0	19–26	15–75	0,1	0,1
Теплый	21,0–22,9	25,1–28,0	20–29	15–75	0,1	0,2

Исходя из требований, рассматриваемых в данном разделе нормативных документов [3, 4], в помещении поддерживается температура равная 19–20 С°, при относительной влажности в 55–58%. В помещении функционирует система принудительной вентиляции. Проводится ежедневная влажная уборка и систематическое проветривание помещения. В зимнее время в помещении предусмотрена система водяного отопления со встроенными нагревательными элементами и терморегуляторами.

5.3.3 Освещенность

Выделяют следующие виды освещения, которые можно использовать при разработке приложения и просмотре видео-контента:

1. естественное;
2. искусственное;
3. совмещенное.

Первый тип освещения наиболее благоприятен для человека с физиологической точки зрения. Однако оно имеет один существенный

недостаток – непостоянство в течение светового дня в зависимости от состояния атмосферы. В помещения естественное освещение попадает через боковые проемы (окна). Следовательно, используется совмещенное освещение (солнечное и искусственное), что соответствует требованиям санитарно-гигиенических норм.

По нормам освещенности СНиП 23-05-95 [5] работа разработчика относится к четвертому разряду зрительной работы средней точности. Рекомендуемая освещенность для данного разряда – в пределах 400 лк.

В помещении, в котором выполнялась разработка, для искусственного освещения используются газоразрядные люминесцентные лампы низкого давления типа ЛБ-40: 2 светильника по 4 лампы в каждом. В нем отсутствуют яркие бликовые поверхности, которые могут оказывать негативное влияние на здоровье и, как следствие, трудоспособность сотрудников. Раз в месяц проводится мытье окон, что позволяет в полном объеме проникать в помещение естественному свету. Однако при его избытке, в том числе при воздействии на сотрудника прямых солнечных лучей существует возможность его ограничения при помощи жалюзи. Освещенность на поверхности стола разработчиков в зоне сосредоточения рабочей документации, а также освещенность экранов видеомониторов примерно равна 300 лк. Светильники общего освещения в зоне углов излучения 50-90° с вертикалью в продольной и поперечной плоскостях не превышает 200 кд/м².

Таким образом данные показатели соответствуют нормам СНиП 23-05-95 [5].

5.3.4 Монотонный режим работы

При работе с персональным компьютером или телевизором существенных факторов, оказывающим негативное влияние на психоэмоциональное состояние разработчика или пользователя, является очень большой по объему информационный поток, который необходимо воспринимать. Поэтому, для снижения воздействия данного вредного производственного фактора, необходимо руководствоваться нормативами

СанПиН 2.2.2/2.4.1340-03. Его соблюдение позволяет увеличить производительность труда и существенно замедлить скорость возникновения профессиональных болезней.

При организации работы с персональным компьютером необходимо делить ее на 3 группы:

- группа А – считывание информации с экрана по предварительному запросу;
- группа Б – ввод информации;
- группа В – режим диалога с персональным компьютером.

Труд разработчика приложения относится к группам А и Б. В свою очередь, деятельность потребителя видео-контента – к группе В. Чтобы добиться снижения вредного воздействия данного фактора необходимо предусматривать перерывы для каждой группы работ – таблица 5.4.

Таблица 5.4 – Суммарное время регламентированных перерывов в зависимости от продолжительности работы, вида категории трудовой деятельности с ПЭВМ [5]

Категория работ с ПЭВМ	Уровень нагрузки за рабочую смену при видах работ с ПЭВМ			Суммарное время регламентированных перерывов при 8-часовой смене, мин.
	группа А, количество знаков	группа Б, количество знаков	группа В, ч	
I	до 20 000	до 15 000	до 2	50
II	до 40 000	до 30 000	до 4	70
III	до 60 000	до 40 000	до 6	90

5.3.5 Пожарная безопасность

Пожары приводят не только к прямым экономическим потерям, но и к получению упущенной выгоды. Полная потеря информации при пожаре, большие трудности ее восстановления могут существенно поколебать действующую бизнес-модель.

Основные источники воспламенения: неисправная электропроводка, короткое замыкание или пробой изоляции; использование поврежденных (неисправных) электроприборов; попадание молнии в здание; безответственное обращение с огнем и несоблюдение мер пожарной безопасности [6].

5.4 Рекомендации по минимизации влияния вредных и опасных факторов производства

5.4.1 Рекомендации по защите от электрического тока

Для защиты от поражения электрическим током все токоведущие части должны быть изолированы от любых видов нерегламентированного прикосновений. Необходимо заземление корпусов устройств. Оно должно быть выполнено изолированным медным проводом с сечением 1,5 мм². В свою очередь этот провод должен быть присоединен к общей шине заземления с общим сечением 5,4 мм². Общая шина присоединяется к заземлению с сопротивлением не более 4 Ом. Электропитание устройства – от силового щита с использованием автоматического предохранителя. Последний должен срабатывать при коротком замыкании нагрузки.

При эксплуатации персональных компьютеров и других ЭВМ нельзя допускать:

- Включение ЭВМ при неисправной защите электропитания;
- При поданном напряжении электросети запрещается подключать и отключать разъемы кабелей электропитания и блоков вентиляции;
- Заменять съемные элементы, находящиеся под напряжением;
- снимать щиты, закрывающие доступ к токоведущим частям;

При правильной эксплуатации электроприборов и использовании соответствующих средств защиты существенно снижается риск поражения электрическим током. Необходимо соблюдать следующие меры безопасности:

- Разработчики и иной персонал должны пройти инструктаж по технике безопасности;
- Осуществлять профилактические проверки исправности электроприборов.

5.4.2 Рекомендации по улучшению микроклимата

Для оздоровление воздушной среды в помещении разработки необходим комплекс мероприятий, к которым могут относиться более рациональная организация систем кондиционирования и вентиляции воздуха, а также отопление помещений. Осуществление вентиляции осуществлять двумя возможными способами: искусственным и естественным. В зимнее время года возможно предусмотреть умную систему отопления, устанавливающую комфортную температуру воздуха.

Желательно, чтобы помещение с персональными компьютерами и другими видами ЭВМ не должен иметь объем меньше 20 м³/человека [7].

Чтобы обеспечить комфортные условия труда необходимо рационально проводить работы в зависимости от времени года и времени суток; чередовать труд и отдых.

5.4.3 Рекомендации по минимизации влияния освещения

Для обеспечения требуемого уровня освещения в помещении используется лампы дневного освещения, равномерно распределенные по всему потолку офиса. Для освещения помещения выбраны наиболее широко применяемые лампы типа ЛБ.

5.4.4 Меры по обеспечению пожарной безопасности

Главным условием обеспечения пожарной безопасности является предупреждение и ликвидация возможных источников воспламенения [6]. В помещении необходимо соблюдать следующие меры:

- Ограничить количества горючих и легко воспламеняющихся веществ;
- Устранить потенциальные источники возгорания (электрические искры, нагрев оборудования);
- Наличие средств пожаротушения, в том числе пожарной сигнализации и содержание их в исправном состоянии;
- Содержание электрооборудования в исправном состоянии;
- Курение в только специально отведенных для этого местах;
- Пути и проходы эвакуации сотрудников и посетителей должны содержаться в свободном состоянии;
- Плановые проведения инструктажей по пожарной безопасности;
- Назначить ответственного за пожарную безопасность.

5.5 Экологическая безопасность

5.5.1 Анализ воздействия на окружающую среду

Для защиты окружающей среды от вредного воздействия выбросов необходимо осуществить полный переход к энергосберегающим и малоотходным (в идеале безотходным) производствам. При разработке программного обеспечения возникает необходимость в утилизации отходов, создаваемых сотрудниками. В первую очередь это бумажные отходы – макулатура, пластиковые отходы, неисправные детали персональных компьютеров и других видов ЭВМ.

5.5.2 Рекомендации по минимизации влияния на окружающую среду

Автор рекомендует накапливать бумажные отходы и передавать их в пункты приема макулатуры для дальнейшей переработки [8]. Пластиковые бутылки складывать в специально предназначенные контейнеры, которые

активно появляются в Томске. Неисправные комплектующие персональных компьютеров и других ЭВМ, а также люминесцентные лампы [9] возможно передавать организациям, занимающимся переработкой отходов [10] или продавать на вторичном рынке.

Дополнительным методом снижения отходов является увеличение доли электронного документооборота.

5.6 Безопасность в чрезвычайных ситуациях

5.6.1 Перечень возможных чрезвычайных ситуаций на объекте

Представим список чрезвычайных ситуаций, которые могут возникнуть при работе в офисе:

- В зависимости от намерения: преднамеренные и непреднамеренные;
- По характеру происхождения:
 - a. Техногенные: пожары, взрывы, нарушение целостности помещений, повреждение или выход из строя систем жизнеобеспечения. Данные причины также могут возникнуть под воздействием сил природы.
 - b. Экологические – это непредвиденные или аномальные изменения состояния окружающей природной среды (ураганы, ливни, паводки, наводнения и паводки). Нарушения в экологической системе может возникать по антропогенным причинам.
 - c. Биологические – эпидемии различных болезней;
 - d. комбинированные.
- В зависимости от скорости развития: плавные, скоротечные, взрывные, внезапные.
- По масштабам последствий: территориальные, местные и локальные.
- По возможности предотвращения: предотвращаемые и неизбежные;

Наиболее распространенной формой чрезвычайной ситуации в офисном здании является возникновение пожара.

5.6.2 Меры по ликвидации чрезвычайных ситуаций и последствий

Пожарная безопасность является одной из приоритетных задач для сотрудников Института кибернетики. В каждом помещении имеется в наличии система защиты, регламентированная законом и нормативными документами.

Сотрудники перед уходом из помещения проводят его осмотр, закрывают окна, и убеждаются в отсутствии источников возможного возгорания, отключают освещение. По внутреннему регламенту один раз в три года проводятся замеры сопротивления изоляции токоведущих частей оборудования. Один раз в полгода сотрудники проходят инструктаж по противопожарной безопасности. При обнаружении признаков горения сотрудники обязаны немедленно вызвать пожарную службу и прекратить работу и, по возможности, эвакуироваться из здания. В случае наличия возможности к тушению пожара своими силами сотрудники должны применять ручные углекислотные огнетушители (типа ОУ-2, ОУ-5). Данные огнетушители находятся в помещениях офиса и предназначены для тушения начальных возгораний. Огнетушители постоянно содержатся в исправном состоянии и готовы к использованию. [11].

5.7 Правовые и организационные вопросы обеспечения безопасности

5.7.1 Психофизиологические факторы

К таким факторам относятся чрезмерное умственное напряжение, статические физические перегрузки опорно-двигательной системы человека, вызванные монотонностью труда, а также перегрузки эмоционального характера. Работнику, в течение каждого рабочего дня необходимо предоставлять перерыв для приема пищи и отдыха. Продолжительность перерыва 30-120 минут. Конкретная величина перерыва устанавливается либо правилами трудового распорядка, либо индивидуальной договоренностью разработчика с руководством. Чтобы сохранить на должном уровне производительность труда и избежать преждевременной утомляемости

разработчика рекомендуется чередовать типы работ и включать такие типы, в которых персональный компьютер не используется вовсе. Также во время перерывов рекомендуется выполнять комплекс физических упражнений.

5.7.2 Организационные мероприятия при компоновке рабочей зоны

Рабочие места разработчиков, работающих на персональных компьютерах, рекомендуется размещать изолированно друг от друга, так как данный вид работы требует высокой концентрации внимания и значительного умственного напряжения. Расположение экрана монитора должно быть на расстоянии 600 - 700 мм от глаз пользователя. Рекомендуется не приближаться ближе чем 500 мм.

Ввиду сидячего характера работы разработчика важнейшим элементом его рабочего места становится кресло. Его конструкция должна поддерживать рациональную рабочую позу при работе на персональном компьютере, а также позволять изменять позу, чтобы снизить статическое напряжения мышц шейно-плечевой области и спины.

Немаловажным фактором является учет требований к размещению средств отображения информации (Таблица 5.5Таблица).

Таблица 5.5 – Расположение средств отображения информации

Тип средств отображения информации	Угол, градусы	
	В горизонтальной плоскости	В вертикальной плоскости
Редко используемые	±60	±60
Часто используемые	±30	±30
Очень часто используемые	±15	± 15

Заключение

Развитие современных вычислительных технологий в настоящее время позволяет внедрять новые и усовершенствовать существующие методы компьютерного зрения в различных сферах человеческой деятельности. Алгоритмы обработки видео и изображений находят свое применение в таких задачах как поиск в коллекциях фильмов и специальных видеофайлах, анализ медицинских снимков, распознавание объектов на спутниковых снимках, а также организация поиска в базах видео данных в сети и на локальном компьютере. Разработанная программа представляет собой средство для поиска видеофрагментов, содержащих заданный образец текстуры. В основе приложения лежит алгоритм выделения текстурных признаков с помощью энергетических карт Лавса. Заметим, что описание текстуры может использоваться в качестве ключа при поиске в видео библиотеках. При этом для обеспечения полноты ключа также предложен модифицированный метод распознавания цветных текстур.

Многообразие методов выделения характерных признаков позволяет постоянно улучшать существующие системы, в основе которых лежит представление графической информации с помощью атрибутов формы, цвета, текстуры, расположения объекта в пространстве. В случае обработки видеофайлов возникает ряд других направлений – исследование оптического потока, который служит для задания уравнений движения объектов на кадре, поиск ключевых кадров и границ сцен. Последнее является важным предварительным этапом, поскольку, наряду с высоким качеством результата, необходимо также обеспечить производительность системы. Использование сложных математических моделей описания видеофайлов зачастую негативно сказывается на скорости обработки, и, несмотря на высокий уровень качества, подобные методы редко можно встретить при решении задач, требующих работы с большими объемами данных. В то же время простые алгоритмы не всегда отвечают требованиям точности поиска. Данный фактор объясняет

отсутствие наилучшего алгоритма выделения характерных признаков и их сравнения.

Исследования предложенного метода показали, что на точность результата значительно влияет разрешение кадра в совокупности со значением входного параметра размера окрестности W . В реальных системах иногда невозможно предоставить видеофайлы высокого разрешения из-за трудностей хранения большого объема информации. Однако для реализованного алгоритма экспериментально было доказано, что качество исходного изображения не влияет напрямую на эффективность обработки. Основным требованием является соответствие разрешения образцов текстур и кадра. Таким образом, система способна обеспечить удовлетворительные результаты даже при небольших размерах исходных данных, избегая при этом негативных последствий потери качества и наличия шума.

В целях дальнейшего улучшения алгоритма предлагается разработка метода нормализации энергетических карт Лавса, для того чтобы избавиться от влияния несоответствия размеров текстур на образце и кадре. Кроме того, одним из основных направлений исследований должно стать дополнение информации о характерных признаках изображения. Возможно применение более сложных моделей описания цвета, например, построение взвешенных гистограмм распределения интенсивностей или кластеризация в различных цветовых пространствах. В основе другого пути модификации метода лежит использование более совершенных метрик для оценки разницы между характеристиками образца и обрабатываемого кадра.

Список используемых источников

1. Квадродеревья и октодеревья [Электронный ресурс] / Лаборатория Обработки Изображений ИВМиМГ СО РАН. – Электрон. дан. – Новосибирск, 2014. – Режим доступа: <http://loi.sccc.ru/gis/QuadTree/QuadTree.html#quadtrees>. – Загл. с экрана.
2. Форсайт Д., Понс Ж. Компьютерное зрение. Современный подход. / Пер. с англ. – М.: Издательский дом «Вильямс», 2004. – 928 с.
3. Фисенко В. Т., Фисенко Т.Ю. Компьютерная обработка и распознавание изображений: Учеб. пособие. – СПб: СПбГУ ИТМО, 2008. – 192 с.
4. Сравнение гистограмм [Электронный ресурс] / Learning OpenCV. – Электрон. дан. – 2014. – Режим доступа: http://locv.ru/wiki/7.4.1_
5. Сравнение_гистограмм. – Загл. с экрана.
6. Примеры использования гистограмм [Электронный ресурс] / Learning OpenCV. – Электрон. дан. – 2014. – Режим доступа: http://locv.ru/wiki/7.4.2_Примеры_использования_гистограмм. – Загл. с экрана.
7. Шапиро Л., Стокман Дж. Компьютерное зрение. / Пер. с англ. – М.: БИНОМ. Лаборатория знаний, 2006. – 752 с.
8. Грузман И. С., Киричук В. С. Цифровая обработка изображений в информационных системах: Учеб. пособие. – Новосибирск: Изд-во НГТУ, 2002. – 352 с.
9. Оглуздина Ю. А. Алгоритмы сегментации изображений // Молодежный научно-технический вестник : электронный журнал. – Москва, 2013.
10. Интерактивная сегментация изображений на основе MRF и алгоритм TRW [Электронный ресурс] / MachineLearning.ru. – Электрон. дан. – [2007 – 2014]. – Режим доступа: http://www.machinelearning.ru/wiki/images/5/5f/Romovpa_smais_trw.pdf. – Загл. с экрана.
11. Сергей Двоенко, Шанг Динь. Распознавание растровых текстурных изображений на основе параметрических ациклических марковских моделей

- // The 22nd International Conference on Computer Graphics and Vision. – Moscow, October 01 – 05, 2012. – p.138 – 143.
12. Main Page [Электронный ресурс] / Emgu CV. – Электрон. дан. – [2009 – 2014]. – Режим доступа: http://www.emgu.com/wiki/index.php/Main_Page. – Загл. с экрана.
 13. Гаврикова Н. А., Финансовый менеджмент, ресурсоэффективность и ресурсосбережение: учебно-методическое пособие / Н. А. Гаврикова, Л. Р. Тухватулина, И. Г. Видяев, Г. Н. Серикова, Н. В. Шаповалова; Томский политехнический университет. – Томск: Изд-во Томского политехнического университета, 2014. – 73 с.
 14. Основы управления проектами [Электронный ресурс] / НОУ «ИНТУИТ». – Электрон. дан. – Москва, [2003 – 2014]. – Режим доступа: <http://www.intuit.ru/studies/courses/2194/272/info>. – Загл. с экрана.
 15. Качество программного обеспечения [Электронный ресурс] : Материал из Википедии — свободной энциклопедии : Версия 63298119, сохранённая в 06:54 UTC 27 мая 2014 / Авторы Википедии // Википедия, свободная энциклопедия. — Электрон. дан. — Сан-Франциско: Фонд Викимедиа, 2014. — Режим доступа: <http://ru.wikipedia.org/?oldid=63298119>.
 16. ГОСТ 12.0.003-74 «Система стандартов безопасности труда. Опасные и вредные производственные факторы. Классификация» (с измен. № 1, октябрь 1978 г., переиздание 1999 г.)
 17. СанПиН 2.2.2/2.4.1340-03 «Гигиенические требования к персональным электронно-вычислительным машинам и организации работ»
 18. СанПиН 2.2.4.548-96 «Гигиенические требования к микроклимату производственных помещений»
 19. СП 52.13330.2011 «Естественное и искусственное освещение. Актуализированная редакция СНиП 23-05-95»
 20. СН 2.2.4/2.1.8.562-96 «Шум на рабочих местах, в помещениях жилых, общественных зданий и на территории жилой застройки»

21. Охрана труда в вычислительных центрах: Учеб. для студентов сред. спец. учеб. заведений по специальности «Программирование для электронно-вычислительных машин и автоматизированных систем» / Ю. Г. Сибаров, Н. Н. Сколотнев, В. К. Васин, В. Н. Нагинаев. – М. : Машиностроение, 1990. – 192 с.
22. Ефремова О. С. Требования охраны труда при работе на персональных электронно-вычислительных машинах. – 2-е изд., перераб. и доп. – М. : Издательство «Альфа-Пресс», 2008. – 176 с.
23. Назаренко О. Б. Безопасность жизнедеятельности: учебное пособие / О. Б. Назаренко, Ю. А. Амелькович; Томский политехнический университет. – 3-е изд., перераб. и доп. – Томск: Изд-во Томского политехнического университета, 2013. – 178 с.
24. Охрана труда при работе с ПК – микроклимат [Электронный ресурс] / Клуб инженеров по охране труда. – Электрон. дан. – Москва, 2014. – Режим доступа: <http://www.dvkuot.ru/index.php/otpk/106-klimat>. – Загл. с экрана.
25. Беляков Г. И. Безопасность жизнедеятельности на производстве (охрана труда): Учебник для вузов. – СПб. : Издательство «Лань», 2006. – 512 с.
26. Последствия воздействия вредных факторов ПЭВМ на пользователя [Электронный ресурс] / Научная библиотека Новосибирского государственного педагогического университета. – Электрон. дан. – Новосибирск, [2000 – 2014]. – Режим доступа: <http://lib.nspu.ru/umk/18f47bea37cb10dc/t8/ch2.html>. – Загл. с экрана.
27. Белов С. В. Безопасность жизнедеятельности и защита окружающей среды (техносферная безопасность): учебник / С. В. Белов. – 2-е изд., испр. и доп. – М.: Издательство Юрайт, 2011. – 680 с.
28. Персональный компьютер: физические факторы воздействия и здоровье пользователя / Ю. Г. Григорьев, О. А. Григорьев // Энергия . – 15/07/1999 . – N 7 . – 29-33 .

29. Энергосбережение в компьютерном мире // HWP. [1998–2012]. Дата обновления: 12.02.2008. URL: http://www.hwp.ru/articles/Energoberezhnie_v_kompyuternom_mire_CHast_1___osnovnie_tendentsii/?SHOWALL_1=1 (дата обращения: 09.11.2012).
30. Безопасность жизнедеятельности. Безопасность технологических процессов и производств (Охрана труда): Учеб. пособие для вузов / П. П. Кукин, В. Л. Лапин, Н. Л. Пономарев и др. – 5-е изд., стер. М.: Высш. шк., 2009 – 335 с.
31. Организация рабочего места при работе за компьютером [Электронный ресурс] / Энциклопедия экономиста. – Электрон. дан. – Москва, 2014. – Режим доступа : <http://www.grandars.ru/shkola/bezopasnost-zhiznedeyatelnosti/organizaciya-raboty-za-kompyuterom.html>. – Загл. с экрана.
32. Дашковский А. Г. Безопасность жизнедеятельности. Защита населения в чрезвычайных ситуациях: учебное пособие / А. Г. Дашковский, И. Г. Романцов – Томск: Издательство Томского политехнического университета, 2008 – 193 с.
33. Histogram Equalization [Электронный ресурс] / Computer Image Processing and Analysis. – Электрон. дан. – 2014. – Режим доступа: http://fourier.eng.hmc.edu/e161/lectures/contrast_transform/node2.html. – Загл. с экрана.
34. Histograms – 2: Histogram Equalization [Электронный ресурс] / OpenCV 3.0.0-dev documentation. – Электрон. дан. – 2011 – 2014. – Режим доступа: http://docs.opencv.org/trunk/doc/py_tutorials/py_imgproc/py_histograms/py_histogram_equalization/py_histogram_equalization.html. – Загл. с экрана.
35. Block-based motion field segmentation for video coding [Электронный ресурс] / Science direct, Elsevier B.V. – Электрон. дан. – 2014. – Режим доступа: <http://www.sciencedirect.com/science/article/pii/S1047320305000271>. – Загл. с экрана.

36. A novel block-based video segmentation algorithm [Электронный ресурс] / SearchGate – Электрон. дан. – 2008 – 2014. – Режим доступа:http://www.researchgate.net/publication/232647098_A_NOVEL_BLOCK-BASED_VIDEO_SEGMENTATION_ALGORITHM. – Загл. с экрана.
37. Image moment [Электронный ресурс] / Wikipedia, The Free Encyclopedia – Электрон. дан. – San Francisco: Wikimedia Fond, 2014. – Режим доступа: http://en.wikipedia.org/w/index.php?title=Image_moment&oldid=607144610
38. Parker, J. R. Algorithms for Image Processing and Computer Vision, Second Edition / J. R. Parker // Wiley Publishing, Inc., 2010 – 504 с.
39. Comparison of Automatic Shot Boundary Detection Algorithms [Электронный ресурс] / Prof. Dr. Rainer Lienhart – Электрон. дан. – Augsburg, 2014. – Режим доступа: http://www.videoanalysis.org/Prof._Dr._Rainer_Lienhart/Publications_files/spie99.pdf

Приложение А

Раздел 6 Video stream segmentation

Студент:

Группа	ФИО	Подпись	Дата
8BM5B	Б.С. Бокарев		

Консультант кафедры ПИ:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент каф. ПИ	Е.С. Чердынцев	к.т.н.		

Консультант – лингвист кафедры ИЯИК:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
зав. каф. ИЯ	Т.В. Сидоренко	к.п.н.		

6 Video stream segmentation

The amount of information in video archives and subject-oriented libraries has the tendency to constant increase. It leads to a problem of fast search of necessary information contained in the archives. In order to avoid processing of each frame, the methods described in this section need to be applied as the first step of video files processing. Video stream contains shorter shots which are basically sets of single pictures. Each of those sets can be described using some features. The main goal of this stage is to find boundaries for the sets. There are also factors complicating this task.

The boundaries may be several types. The simplest situation is cut, when there is a combination of absolutely different sets of frames. Dissolve is smooth transition between shots, when two pictures overlay each other for a short period of time. Other type of boundaries is fade-out/fade-in characterizing with slowly changing intensity of an image.

The detection of such boundaries is an important step of video indexing. In addition, it is necessary to use an algorithm which is effective in searching a shot with great difference to the previous one. Some of these algorithms are based on pixel comparison between two frames, which requires a lot of calculations. Histogram methods are more effective than the first group methods, because they are resistant to small camera movements. Block methods are quite similar to pixel-based approach, but they use several square or rectangular blocks to present an image. The concept of optical flow allows calculating equations of motion. It is also possible and reasonable to combine advantages of few methods shortly described above for some tasks.

6.1 Pixel-based methods

The simplest implementation of the algorithm defines global difference between pixel intensity for two adjacent frames at the moment t and $t + 1$ [34]:

$$d(t,t+1) = \frac{|\sum_{x=1}^N \sum_{y=1}^M I_t(x,y) - \sum_{x=1}^N \sum_{y=1}^M I_{t+1}(x,y)|}{N \cdot M}, \quad (6.1)$$

where N is width of the frame, pixels;

M is height of the frame, pixels;

$I_t(x,y)$ is the intensity of point (x,y) at the moment t ;

The difference may be also calculated between pixels with the same location on two frames. Formula (6.2) is used for greyscale images, while (6.3) allows to compute the value for color images.

$$d(t,t+1) = \frac{\sum_{x=1}^N \sum_{y=1}^M |I_t(x,y) - I_{t+1}(x,y)|}{N \cdot M}, \quad (6.2)$$

$$d(t,t+1) = \frac{\sum_{x=1}^N \sum_{y=1}^M \sum_c |I_t(x,y,c) - I_{t+1}(x,y,c)|}{N \cdot M}, \quad (6.3)$$

where N is width of the frame, pixels;

M is height of the frame, pixels;

$I_t(x,y)$ is the intensity of a point (x,y) at the moment t ;

$I_t(x,y,c)$ is the intensity of color component c at the moment t ;

If the value of the difference $d(t,t+1)$ is higher than some threshold amount δ .

Future improvement of the method may be reached by using statistic values, for example normalized energy of differences (6.4) and normalized sum of a difference module (6.5) [35].

$$d(t,t+1) = \frac{\sum_{x=1}^N \sum_{y=1}^M (I_t(x,y) - I_{t+1}(x,y))^2}{(\sum_{x=1}^N \sum_{y=1}^M (I_t(x,y))^2)(\sum_{x=1}^N \sum_{y=1}^M (I_{t+1}(x,y))^2)} \quad (6.4)$$

$$d(t,t+1) = \frac{\sum_{x=1}^N \sum_{y=1}^M |I_t(x,y) - I_{t+1}(x,y)|}{\sum_{x=1}^N \sum_{y=1}^M I_t(x,y) + \sum_{x=1}^N \sum_{y=1}^M I_{t+1}(x,y)} \quad (6.5)$$

The main disadvantage of this approach is its inability to recognize the boundary between two sequences with small differences in a wide region and wide differences in small region. For instance, if there is a fast changing of a frame contents, this method will find a false boundary. This makes the algorithm very sensitive to camera and object movements. It often happens while processing frames with panorama views, or scaling.

6.2 Histogram analysis methods

While the methods of the group described in the previous section are local, the histogram analysis is based on using global characteristics of an image. It enables to make these algorithms more resistant to camera and objecting movements. Moreover, histograms are invariant to small rotation and have weak dependence on scaling. But sometimes two absolutely different images have almost identical histograms, because they do not account for location of pixels.

Histogram methods have a lot of variants. The simplest one calculates the histogram difference of two frames. More complex algorithms account for weight coefficients, cross product, and special measure of identity.

Formula (6.6) shows the concept of the simplest approach.

$$d(t,t+1) = \sum_{v=0}^V |H(I_{t+1}, v) - H(I_t, v)| , \quad (6.6)$$

where $d(t,t+1)$ – the difference between intensity histograms H ;

V – the amount of histogram bins.

For color images formula (6.7) needs to be used.

$$d(t,t+1) = \max_{k \in \{R,G,B\}} \sum_{v=0}^V |H(I_{t+1}, c_k, v) - H(I_t, c_k, v)| , \quad (6.7)$$

where c_k – color component.

Some color components might be more important than others, therefore it is not always necessary to use all three components of color space. In formula (6.8) the idea of histogram method for HSV color space is shown. Here only 2 components are used – hue and saturation.

$$d(t,t+1) = \sum_{k=1}^2 \sum_{v=0}^V |H(I_{t+1}, c_k, v) - H(I_t, c_k, v)| \quad (6.7)$$

Calculating the histogram difference for only hue-component is shown below in formula (6.8).

$$d(t,t+1) = \frac{\sum_{v=0}^V |H(I_{t+1}, c_k, v) - H(I_t, c_k, v)|}{\sum_{v=0}^V H(I_t, c_k, v)} \quad (6.8)$$

As it is important to provide good speed of processing, reduced RGB color space is used. As the result histograms contain 64 bins, for example (6.9).

$$d(t,t+1) = \frac{\sum_{v=0}^{63} |H_{64}(I_{t+1}, c_k, v) - H_{64}(I_t, c_k, v)|}{\sum_{v=0}^{63} H_{64}(I_t, c_k, v)} \quad (6.9)$$

The boundary between frames sequences can be defined by calculating difference between histograms with weight coefficients which allows considering the importance of a certain color component (6.10).

$$d(t,t+1) = \sum_{k=1}^3 \sum_{v=0}^V \frac{I_t(c_k)}{I_{\text{mean}}} |H(I_t, c_k, v) - H(I_{t+1}, c_k, v)|, \quad (6.9)$$

where $I_t(c_k)$ – intensity of k-color component of t-frame;

I_{mean} – average intensity for t-frame over all color components.

Other approach for calculating weight coefficients is shown in formula (6.10).

$$d(t,t+1) = \sum_{k=1}^3 \left(\sum_{v=0}^V H(I_t, c_k, v) v - \sum_{v=0}^V H(I_{t+1}, c_k, v) v \right)^2, \quad (6.10)$$

where d is the difference between average histogram colors.

6.3 Block-based methods

These methods intend to increase segmentation quality and capacity, since they consider space inhomogeneity of images, in contrast to pixel- and histogram-based algorithms. They also are not sensitive to noises, and object and camera movements. The main concept is to divide an image to blocks and apply methods described in previous sections.

Usually, blocks with same location in two adjacent frames are compared to each other. Measure of identity is based on average value and dispersion, calculated for all image. The identity metric for a block b of two frames at moments t and t+1 can be defined with formula (6.11) [36].

$$d(I_t, I_{t+1}, b) = \frac{\left[(\sigma_{t+1,b}^2 + \sigma_{t,b}^2) + \left((\mu_{t+1,b} - \mu_{t,b}) / 2 \right)^2 \right]^2}{\sigma_{t+1,b}^2 \sigma_{t,b}^2}, \quad (6.11)$$

where $\mu_{t,b}$ is average value of pixels in block b at the moment t;

$\sigma_{t,b}^2$ is dispersion of pixels in block b at the moment t.

Next step is comparison of $d(I_t, I_{t+1}, b)$ with threshold value (6.12).

$$d_D(I_t, I_{t+1}, b) = \begin{cases} 1, & d(I_t, I_{t+1}, b) > \delta_D \\ 0, & \text{otherwise} \end{cases} \quad (6.12)$$

The boundary between two frames is registered if the expression (6.13) is true.

$$\sum_{b=1}^B c_b d_D(I_t, I_{t+1}, b) > \delta, \quad (6.13)$$

where B is total amount of blocks;

c_b is coefficients used for giving the importance to some blocks.

Other way to compute the identity metric is to use Yakimovsky rate (6.14) [37].

$$d(I_t, I_{t+1}, b) = \frac{(\sigma_{\{t,t+1\},b}^2)^2}{\sigma_{t+1,b}^2 \sigma_{t,b}^2}, \quad (6.14)$$

where $\sigma_{\{t,t+1\},b}^2$ is pixel intensity dispersion of block b at the moments t and t+1.

Some modifications offer to combine histogram difference with identity metric [36]. In this case, method contains 2 steps including 3 threshold values definition. First step is calculating the histogram difference $d(I_t, I_{t+1})$, which need to be compared with the threshold for the purpose of false boundaries detection. If it is less than threshold, it is compared with the second threshold value. Second step includes calculating the identity metric. It starts when $d(I_t, I_{t+1})$ more than the second threshold. Each of 2 frames is divided to 64 blocks, and 16 central blocks are used for future processing. The identity metric is calculated for a block b at the moment t+1 and blocks b' at the moment t. b' is the nearest neighbors of block b. Next step is

choosing the smallest one, and finding the average value for such 16 minimums. This value is compared to the third threshold, and if it is higher, the boundary will be registered.

6.4 Feature extraction

Described above methods are based on quite simple image characteristics. Here more complex features will be discussed.

Invariant moments combined with histogram intersection are often used for frame boundaries detection [39]. The moment m of an image I_t is defined with the formula (6.15).

$$m_{p,q} = \sum_{x=1}^N \sum_{y=1}^M x^p y^q I_t, \quad p, q = 0, 1, 2, \dots \quad (6.15)$$

This means that if the pixel intensity is taken like the pixel mass, the whole image mass will be m_{00} , inertness moments at x and y directions will be $m_{2,0}$ and $m_{0,2}$ correspondingly.

The advantage of using of invariant moments consists in their permanency when scaling, rotating, and moving.

The frame model at the moment t is presented by a vector consisting of invariant moments Φ_i (6.16).

$$\vec{\sigma}(t) = \{\Phi_1, \Phi_2, \dots, \Phi_7\} \quad (6.16)$$

For calculating the distance d_1 between two frames the Euclidean metric is often used (6.17).

$$d_1(t_1, t_2) = |\vec{\sigma}(t_1) - \vec{\sigma}(t_2)| \quad (6.17)$$

The histogram intersection is defined by formula (6.18).

$$d_1(t, t+1) = \frac{\sum_{v=0}^V \min(H(I_{t+1}, v), H(I_t, v))}{\sum_{v=0}^V H(I_t, v)}, \quad (6.18)$$

where $H(I_t, v)$ is the histogram for an image I at the moment t ;

V – the amount of histogram bins.

After the values of d_1 and d_2 are calculated, their combination is conducted using the formula (6.19).

$$\Omega(\zeta) = \begin{cases} 3, & \zeta < \delta_1 \\ 2, & \delta_1 < \delta_2 \\ 1, & \delta_2 < \zeta \end{cases}, \quad (6.18)$$

where $\zeta = d_1(t, t + 1)$ and $\Omega_{\text{mom}}(d_1(t, t+1)) \in \{1, 2, 3\}$ for invariant moments;

$\zeta = d_2(t, t + 1)$ and $\Omega_{\text{hist}}(d_2(t, t+1)) \in \{1, 2, 3\}$ for histograms;

δ_1 and δ_2 are threshold values;

$\Omega = 1$ is when two frames are different;

$\Omega = 2$ refers to a boundaryline case;

$\Omega = 3$ means that frames have high level of identity.

Table 6.1 enables to make the final decision on frame identity.

Table 6.1 – Correspondence between Ω_{mom} and Ω_{hist}

Invariant moment	Histogram	Result
3	3	3
3	2	3
3	1	2
2	3	3
2	2	2
2	1	1
1	3	1
1	2	1
1	1	1

Another approach uses brightness level to recognize the boundaries [39]. Its rapid change usually appears on the object edge, or when the orientation of a surface is changing. Edge points with high level of brightness level changing may be used for video sequences segmentation. The main concept of the method is that if new edges appear at the completely different region of the frame in comparison to previously detected edges, it means that the boundary between two video fragments is found.

6.5 Motion-based methods

Since motion is a continual process, it may be used for boundary detection. There are several approaches in this group but generally they all consist in

comparison of characteristics of moving blocks. Usually a match block in frame t+1 will be searched when processing frame t. For this purpose, normalized correlation between blocks and its maximum value need to be calculated. This step is time-consuming, so it might be replaced with a normalized correlation in a frequency domain (6.20) [39].

$$\rho(\xi) = \frac{F^{-1} \{ \hat{x}_1(\omega) \hat{x}_2^*(\omega) \}}{\sqrt{\int |\hat{x}_1(\omega)|^2 d\omega \cdot \int |\hat{x}_2(\omega)|^2 d\omega}}, \quad (6.20)$$

where ξ and ω is a space and frequency coordinates;

$\hat{x}_i(\omega)$ means Fourier transform for block $b_i(\xi)$;

F^{-1} is an inverse Fourier transform operator;

* is a complex conjugation operator.

Before calculating the value (6.20) a high frequency band pass filter is applied to an image which makes the correlation invariant to average intensity changing, because it is not passed by the filter. Normalization of correlation enables the method to be not sensitive to intensity level. Through the use of the value of the correlation coefficient maximum, the location of certain block, when transition from a frame t to a frame t+1, can be defined. The extent of compliance value $p_i = \max(\rho(\xi))$ lies in the range from 0 (absolute difference) to 1 (absolute match). If frames belong to same set, this value will be close to 1.

The extent of compliance for two frames is defined by combination of the block extent of compliance, or μ – the average value of the extent of compliance (6.21).

$$\mu = \frac{\sum_{i=1}^B p_i}{B}, \quad (6.21)$$

where B is a total number of blocks.

Among a set of frames it might be blocks with a very small extent of compliance because of object overlap. To avoid such situations the values of p_i , out of a standard error distance, are not taken into account.

$$\bar{d} = \frac{\sum_{i=1}^{t-1} d_i}{t-1}, \quad (6.22)$$

where the count starts from the last registered set boundary. Then a new set boundary is registered if the expression (6.23) is true.

$$\bar{d} - d_i > \delta, \quad (6.23)$$

where d_i is a total extent of compliance;

δ is a threshold.

6.6 Video keyframe extraction

After the initial video sequence is divided to sets of frames, we need to identify the sets' contents. For this purpose, methods of video keyframes extraction are applied to the sets.

Keyframes are certain frames which characterize its set more clearly. The simplest approach is to take the first or the middle frame in a set as a keyframe. But if there is a complex dynamic motion in a video, a keyframe consisting of only one image is not enough. The number of images in a keyframe depends on a type of video stream.

One of the possible methods of the keyframe subset extraction is based on discrete cosine transform (DCT). Two-dimensional DCT $AC(p,q)$ of an image $I(x,y)$, where p and q are frequency indexes, is defined in formula (6.24) [39].

$$AC(p,q) = \alpha_p \alpha_q \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} I(x,y) \cos \frac{\pi(2x+1)p}{2N} \cos \frac{\pi(2y+1)q}{M}, \quad (6.24)$$

where $0 \leq p \leq N-1$, $0 \leq q \leq M-1$, and

$$\alpha_p = \begin{cases} \frac{1}{\sqrt{N}}, & p=0 \\ \sqrt{2/N}, & 1 \leq p \leq N-1 \end{cases}, \quad (6.25)$$

$$\alpha_q = \begin{cases} \frac{1}{\sqrt{M}}, & q=0 \\ \sqrt{2/M}, & 1 \leq q \leq M-1 \end{cases} \quad (6.26)$$

The value of extent of compliance $k_{opt}(p,q)$ is chosen relying on the criteria (6.27), where $R_i(p,q)$ is optimal coefficients DCT for the i -set.

$$k_{opt}(p,q)=\operatorname{argmin}_k \{|DCT_i(p,q,k) - DCT_{query}(p,q)|\}, \quad (6.27)$$

where $1 \leq k \leq k_d(p,q)$;

$DCT_{query}(p,q)$ are DCT coefficients calculated for an each query.

The distance between the frame and i -set is evaluated by formula (6.28).

$$\operatorname{Dis}(i)=\sum_{p=0}^6 \sum_{q=0}^6 \left(R_i(p,q)-DCT_{query}(p,q)\right)^2, \quad p+q \leq 6 \quad (6.28)$$

When a camera moves, the keyframe should consists of several frames. This task is solved by using oriented graphs with weighted branches. The shortest distance between two graph nodes will be the distance with the least total weight. To avoid redundancy the extent of compliance should be minimalized. Therefore the graph nodes correspond to the video sequence frames and weight coefficients equal to the extent of compliance [7]. The frames with the shortest distance will be the keyframe subset.

Another approach is based on simple geometrical characteristics of video sequences [39]. When transition between frames, three kinds of regions are formed:

- disappeared region which is present on $t-1$ frame and is absent on t frame;
- general region which is present on both t and $t-1$ frames;
- new region which appears on t frame.

For registering new keyframes this method estimates the sizes for each of those regions.

Conclusion

This section described known methods of video data pre-processing. As video is basically the sequences of images, algorithm for image processing can be easily applied to video data. The problem consists in efficiency of some of them, since they may require high productivity. If each frame of video is processed, it will take too

much time to execute some arithmetic calculations. Therefore the initial video sequence is divided to several parts, each of which is presented by the keyframe.

The choice of an algorithm depends on the task. For instance, when processing monitor camera video, it is usually is greyscale and have no motion. More complicated situation appears when initial video set is an artistic movie with complex transitions between frames, scaling, and rotation.

Keyframe extraction algorithms' analysis is very important, since it lets reduce the amount of information significantly and produces direct impact on entire system efficiency.