Федеральное государственное бюджетное учреждение науки
Институт математики и механики им. Н.Н.Красовского
Уральского отделения Российской академии наук

На правах рукописи

КОРНИЛОВ Федор Андреевич

РАЗРАБОТКА МЕТОДОВ РАСПОЗНАВАНИЯ СТРУКТУРНЫХ РАЗЛИЧИЙ ИЗОБРАЖЕНИЙ

05.13.17 — теоретические основы информатики

Диссертация
на соискание ученой степени
кандидата физико-математических наук

Научный руководитель:
КОСТОУСОВ Виктор Борисович
кандидат физ.-мат. наук, с.н.с.

Екатеринбург — 2015

Оглавление

Введение **4**

1 Задача поиска структурных различий изображений **13**

1.1 Основные обозначения 13

1.2 Обзор литературы 14

1.2.1 Геометрическая и радиометрическая коррекция входных изображений . . 14

1.2.2 Поиск существенных различий 17

1.2.3 Пороговая обработка и оценка точности работы 24

1.3 Существующие подходы к определению структуры 26

1.4 Место данной работы в сложившейся системе методов 28

2 Модель структурных различий и задача их локализации **30**

2.1 Определение структуры изображения 30

2.2 Обобщенная яркостная коррекция изображений 34

2.3 Структура изображения и морфологический проектор 36

2.4 Обнаружение и локализация структурных различий 38

2.5 Математическая модель структурных различий и постановка задачи 39

2.6 Решение задачи локализации структурных различий 42

2.7 Выводы по главе 2 56

3 Алгоритмы поиска структурных различий и их реализация **57**

3.1 Общая схема алгоритма поиска структурных различий 58

3.2 Алгоритм поиска структурных различий на основе морфологического проектора 60

3.2.1 Регуляризация морфологического проектора 61

3.2.2 Морфологический проектор для цветных изображений 65

3.3 Алгоритмы поиска структурных различий для отдельных классов функций яркостной коррекции изображений 66

3.3.1 Класс линейных функций преобразования яркости 69
3.3.2 Класс квадратичных функций преобразования яркости 70
3.4 Алгоритм поиска структурных различий на основе глобальной оптимизации . . . 72
3.5 Алгоритм поиска структурных различий без использования яркостной коррекции изображений 75
3.6 Асимптотическая оценка времени работы алгоритмов 79
3.7 Программный комплекс для поиска структурных различий 87
3.8 Использование графических процессоров 91
3.9 Выводы по главе 3 93

4 Исследование алгоритмов поиска структурных различий 94
4.1 Оптимальный порог в случае отсутствия шума 95
4.2 Математическая модель для проведения численного эксперимента 97
4.3 Анализ ROC-кривых 99
4.4 Оценка оптимальных параметров алгоритмов для зашумленных изображений . . 100
4.5 Сравнение результатов работы алгоритмов поиска структурных различий 115
4.6 Исследование влияния рассинхронизации изображений на поиск структурных различий 122
4.7 Выводы по главе 4 124

Заключение 125

Литература 130

Введение

Актуальность темы. Задача поиска существенно изменившихся областей (*структурных различий*) на последовательности изображений является классической для анализа изображений и возникает в различных областях компьютерного зрения, таких как сжатие видеоданных, системы видеонаблюдения, дистанционное зондирование Земли и других [104]. Вне зависимости от специфики конкретной задачи, алгоритмы поиска структурных различий состоят из трех базовых шагов: предварительная обработка входных изображений (геометрическая и радиометрическая коррекция), собственно поиск различий и обработка полученных результатов (оценка площади и/или типа произошедших изменений). Методы непосредственно поиска различий можно разделить на алгебраические (например, вычитание изображений), трансформацию (уменьшение размерности данных, например, анализ главных компонент), классификацию (в частности, текстурную) и другие [91]. Фундаментальное отличие между этими подходами сводится к различному пониманию структуры изображения.

Единого определения структуры изображения не существует, а каждая конкретная вариация данного понятия связана со спецификой рассматриваемой задачи. Однако среди многообразия подходов, предложенных для обработки изображений, следует отметить морфологический анализ изображений как наиболее близкую к указанной задаче область, для которой существует несколько проработанных теорий. Среди них морфологическая теория Серра [109], теория многомасштабного анализа [79] и морфологический анализ изображений Ю.П. Пытьева [39, 40]. В работах последнего развиты методы, инвариантные относительно преобразований, характеризующих влияние условий регистрации изображений. Ключевым понятием данной теории является форма изображения, под которой понимается оператор проектирования в пространстве изображений на множество изображений, полученных из исходного путем применения всех возможных преобразований яркости. Этот оператор проектирования — морфологический проектор — в данной работе является основой алгоритмов поиска структурных различий изображений.

В качестве сферы применения предлагаемых в диссертации методов рассмотрена одна из наиболее важных задач анализа цифровых космических снимков земной поверхности: поиск изменений на двух разновременных изображениях с целью выявления различий в

составе объектов сцены на них. Решение данной задачи необходимо для своевременного обновления топографических карт и постоянного оперативного мониторинга земной поверхности, являющихся, с одной стороны, актуальными вследствие стремительного развития городской и транспортной инфраструктуры, вырубки лесов и других изменений природного или антропогенного характера, а с другой стороны — весьма трудоемких и решающихся в настоящее время в ручном или полуавтоматическом режиме человеком-оператором. Поэтому решение об обновлении, в частности, топографических карт принимается в случае, если на местности произошли существенные изменения в составе объектов. Здесь, в неформальном, содержательном смысле под структурными различиями понимаются существенные изменения местности, при которых объекты интереса появляются, исчезают или изменяют свою форму. При этом изменения освещенности и цвета объектов структурными различиями не считаются (рис. 0.1).

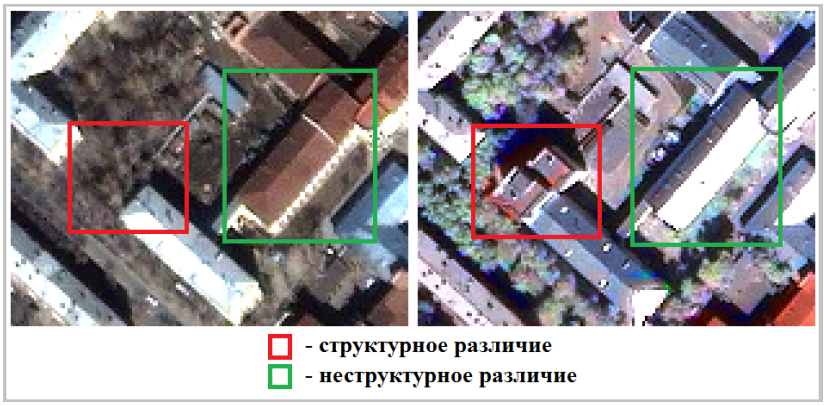

Рис. 0.1. Пример структурного и не структурного различия изображений. В левом прямоугольнике выделен появившийся дом: это — структурное различие. В правом — дом, яркость крыши которого существенно изменилась, но форма осталась прежней: это — не структурное различие.

Важным моментом является то, что необходимо не просто обнаружение и выделение структурных различий, а определение общей площади этих изменений. Такой подход позволяет с помощью среднемасштабных снимков определить степень изменения некоторого участка земной поверхности, после чего принять решение о необходимости заказа крупномасштабных снимков для этой области и обновления соответствующих топографических карт.

Эти соображения делают автоматизацию задачи поиска структурных различий, основанную на дешифрировании аэрокосмических фотоснимков, востребованной в различных геоинформационных системах (ГИС) [44], используемых во многих сферах человеческой деятельности. Основными этапами дешифрирования являются обнаружение, распознавание, определение количественных и качественных характеристик, а также регистрация результатов дешифриро-

вания топографических объектов местности. Несмотря на значительный прогресс в направлении повышения разрешающей способности и скорости получения таких данных, наиболее сложные и трудоемкие этапы дешифрирования (обнаружение и распознавание) в настоящее время выполняются вручную. На современном этапе развития основным путем повышения эффективности дешифрирования материалов аэрокосмической съемки является разработка средств и методов автоматизированного дешифрирования изображений. Трудность автоматизации подобных процессов связана с большой вариативностью изображений объектов земной поверхности и широким спектром решаемых задач [12, 93]. Появляющиеся в связи с этим задачи распознавания и вопросы обнаружения изображений объектов местности все больше привлекают внимание многих исследователей, так как являются одними из наиболее значимых областей применения и развития средств обработки изображений, распознавания образов, компьютерного и технического зрения [47].

С точки зрения требований к входным данным в задаче поиска структурных различий изображений рассматриваются два разновременных цифровых космических снимка одного и того же участка земной поверхности. Предполагается, что они геометрически выровнены и имеют одинаковый размер в пикселях. Геометрическая выравненность означает, что каждый пиксель на обоих снимках имеет одни и те же координаты на местности, т. е. эти снимки одинаково привязаны к местности. Как правило, работа осуществляется с полутоновыми изображениями, однако отдельно будет рассмотрен вариант обработки многоканальных (цветных) изображений.

Степень разработанности темы. Задача поиска существенно изменившихся областей (*change detection*) на последовательности изображений вот уже много лет привлекает внимание исследователей по всему миру: Alves D. S., Bauer M. E., Brondizio E., Bruzzone L., Collins J. B., Congalton R. G., Coppin P. R., Franklin S. E., Fernandez Prieto D., Fung T., Jensen J. R., Huertas A., Lambin E. F., Moran E., Mausel P., Nelson R., Nevatia R., Serpico S., Townshend J. R. G., Tucker C. J., Woodcock C. E., а среди отечественных авторов — Ю. В. Визильтер, С. Ю. Желтов, Л. М. Местецкий, Ю. П. Пытьев, В. В. Сергеев, А. И. Чуличков, А. Ю. Рубис, В. С. Горбацевич. Причиной такой популярности является многообразие прикладных задач, для решения которых требуется знание о произошедших в изучаемой сцене изменениях: это задачи видеонаблюдения (в которых поиск различий является предобработкой с целью уменьшить объем входных данных), видеокодирования, анализ движения, медицинские приложения и анализ космоснимков (мониторинг земной поверхности). Вследствие этого разработано и предложено необычайно широкое многообразие практических алгоритмов, позволяющих решать данную задачу. Одна из самых важных проблем при решении задачи поиска различий связана с определением того, что является этим самым различием, т. е. с необходимостью построения специальных мате-

матических моделей для существенно изменившихся областей. Поэтому поиск различий более правильно называть поиском *существенных* различий, и каждый алгоритм настраивается на специфику конкретных практических требований, иными словами — на то, что считать *структурой* изображения. Существует множество различных определений структуры изображений, среди которых можно выделить следующие три: морфологическая теория Серра, теория многомасштабного анализа и теория морфологического анализа Ю. П. Пытьева. Последняя является теоретической основой данной диссертации, в частности, исследуется предлагаемый в работах Ю. П. Пытьева морфологический проектор, являющийся мощным методом устранения влияния условий регистрации изображений.

Предлагаемая диссертационная работа посвящена исследованию одного класса алгоритмов поиска структурных различий изображений, основанных на предварительном относительном яркостном выравнивании исходных изображений для устранения изменений, связанных со сменой условий их регистрации или наличием несущественных различий, и их последующего вычитания для обнаружения изменившихся областей. Такой подход обладает рядом преимуществ, в частности, гибкость и простота реализации, а также малое время работы. Наличие ряда алгоритмов порождает необходимость в методике их сравнения и исследования. Важную роль здесь играет оценка оптимальных параметров алгоритмов, например, порогового значения, оказывающего значительное влияние на результаты работы. При этом подобная работа должна проходить в рамках ограничений, налагаемых спецификой конкретной прикладной задачи, что делает математическую формализацию задачи, включающую в себя определение структурного (существенного) различия и его математическую модель, не менее актуальной темой для исследования.

Представленная в диссертации задача поиска структурных различий появилась в рамках задач корректировки карт и навигации по геофизическим полям [2,9,10,12–14,32,33,80], которыми занимается Отдел прикладных проблем управления ИММ УрО РАН. Вторым направлением работы отдела является решение важных прикладных задач технического и компьютерного зрения, имеющих перспективы широкого практического применения. Большой опыт, касающийся алгоритмов обработки изображений и технического зрения, накопленный за время решения этих задач, оказал существенное влияние на изложенные в данной работе результаты.

Цель и задачи исследования. Целью работы является создание и исследование алгоритмов поиска структурных различий изображений из класса методов, основанных на предварительном относительном яркостном выравнивании исходных изображений. Для ее достижения в диссертации решаются следующие задачи:

1. Разработка математической формализации задачи поиска структурных различий изоб-

ражений, включающей математическую модель структурных различий и постановку задачи их обнаружения. Исследование распределения ошибок обнаружения и вычисление оптимальных параметров алгоритма поиска различий, построенного на основе морфологического проектора.

2. Создание новых прикладных алгоритмов поиска структурных различий.

3. Разработка методики и проведение вычислительного эксперимента для сравнения и исследования различных алгоритмов.

4. Построение на основе разработанных алгоритмов программного комплекса, решающего задачу оценки изменчивости территорий по данным дистанционного зондирования Земли.

Научная новизна. Научная новизна работы заключается в разработке автором формализации задачи поиска структурных различий изображений, включающей в себя определение структурного различия и его математическую модель. В рамках данной формализации впервые получена формула распределения вероятности яркостей пикселей изображения, представляющего собой разность исходного изображения и его морфологической проекции, а также формула для оптимального порога алгоритма поиска структурных различий, основанного на морфологическом проекторе. Предложена оригинальная формула обобщенной яркостной коррекции изображений, на основе которой разработан ряд новых прикладных алгоритмов, решающих поставленную задачу поиска структурных различий изображений. Для оценки оптимальных параметров алгоритмов и выбора наиболее подходящего для практического применения и разработки программного комплекса предложена модель и схема проведения вычислительного эксперимента.

Теоретическая и практическая значимость работы. *Теоретическая ценность* работы состоит в том, что в ней приведена математическая формализация задачи поиска структурных различий изображений, развивающая теоретическую основу для данного класса задач обработки изображений и позволяющая получать строгие утверждения о работе и оптимальных параметрах алгоритмов поиска структурных различий. *Практическая ценность* работы заключается в выводе формулы оптимального порога алгоритма поиска структурных различий на основе морфологического проектора для разработанной в рамках формализации задачи математической модели. Предложенная формула обобщенной яркостной коррекции изображений позволяет строить новые алгоритмы, настроенные на решение конкретных прикладных задач в ГИС приложениях. Также предложена модель и схема вычислительного эксперимента для эмпирического исследования и сравнения различных алгоритмов поиска структурных различий. Кроме того, в рамках диссертации создан программный комплекс, решающий задачу поиска и анализа степени структурных различий изображений, который был интегрирован в среду визуализации и обработки данных дистанционного зондирования Земли ENVI. Также была разра-

ботана версия программы, использующая высокопроизводительные многоядерные графические процессоры.

Методология и методы исследования. Используются методы теории вероятностей, математической статистики, математической морфологии, обработки изображений, математического моделирования.

Положения, выносимые на защиту. На защиту выносятся следующие новые научные результаты:

1. На основе понятия структуры изображения как набора множеств уровней яркости введена вероятностная модель структурных различий пары изображений и сформулирована задача поиска структурных различий. Для этой модели впервые получена конструктивная формула распределения вероятности яркости точек разностного изображения, представляющего собой разность исходного изображения и его морфологической проекции. С помощью этой формулы построен оптимальный порог исследуемого алгоритма.

2. Разработан ряд новых прикладных алгоритмов, решающих задачу поиска структурных различий изображений, использующих как сканирующее окно, так и глобальную оптимизацию энергетической функции. Также разработан алгоритм, основанный на непосредственном применении введенного определения структурных различий, полезный как для поиска малоразмерных структурных различий, так и для устранения локальной рассинхронизации изображений.

3. Предложена специализированная методика численного эксперимента для сравнения качества работы алгоритмов поиска структурных различий изображений. Эта методика построена на базе параллельных суперкомпьютерных технологий и, в частности, ее программная реализация использует высокопроизводительные многоядерные графические процессоры. С помощью данной методики проведен объемный вычислительный эксперимент, результаты которого позволили оценить оптимальные параметры новых и существующих алгоритмов.

4. Созданный на базе проведенного исследования программный комплекс, решающий задачу оперативной оценки изменчивости топографических объектов местности по разновременным космическим снимкам. Данный комплекс был интегрирован в среду визуализации и обработки данных дистанционного зондирования Земли ENVI и внедрен в практику научно–исследовательского и производственного центра АО “НИиП центр Природа”, (г. Москва).

Степень достоверности результатов. Математические утверждения, полученные в рамках диссертационного исследования, снабжены строгими доказательствами, что обеспечивает достоверность результатов. Теоретические положения подтверждены вычислительными экспериментами по сравнительному анализу предложенных методов.

Апробация работы. Основные результаты докладывались и обсуждались на следующих международных и всероссийских научных конференциях:

— всероссийских и международных молодежных школах–конференциях “Проблемы теоретической и прикладной математики” (Екатеринбург, 2010 — 2013 гг.);

— научно-технической конференции-семинаре “Техническое зрение в системах управления мобильными объектами” (Таруса, 2010 г.);

— научно-технических конференциях “Техническое зрение в системах управления” (Москва, 2012, 2013 гг.);

— десятой международной конференции “Интеллектуализация обработки информации–2014” (о.Крит, Греция, 2014 г.);

— семнадцатой всероссийской конференции с международным участием “Математические методы распознавания образов–2015” (Светлогорск, 2015 г.).

Результаты диссертации опубликованы в 15 научных работах [4, 17–28, 30, 31], из них 2 в изданиях, входящих в список ВАК. Результаты работы внедрены в ПК “ДЕКОС” и модуль для поиска структурных различий на разновременных космических снимках земной поверхности, реализованный в ГИС ENVI (совместно с АО “НИиП центр Природа”).

Структура и объем работы. Диссертация состоит из введения, четырех глав основного содержания, заключения и списка цитируемой литературы. Общий объем работы составляет 140 страниц машинописного текста; диссертация содержит 44 рисунков, 9 таблиц и 122 ссылки на литературные источники.

Содержание работы. Вследствие своей практической важности, задаче поиска различий на последовательности изображений уделяется значительное внимание исследователей по всему миру. В **первой главе, “Задача поиска структурных различий изображений”** приводится обзор существующих методов по данной тематике. Дается общий список операций, необходимых для обнаружения различий, и приводится описание алгоритмов, их выполняющих. Также рассматриваются существующие определения понятия структуры изображения и определяется место предлагаемых в данной работе алгоритмов среди предшественников.

Во **второй главе, “Модель структурных различий и задача их локализации”** вводится формализация задачи поиска структурных различий изображений: рассматривается вопрос определения структуры изображения на основе геометрических и/или яркостных связей между парами пикселей или их группами, и вводится соответствующее определение структурных различий. На основе данных определений конечная задача может быть сформулирована двумя способами: как поиск структурных различий вообще (т. е. определение того, что входные изображения имеют разную структуру), либо как локализация структурных различий в каж-

дой точке изображений. Исходя из этого разделения производится построение математической модели структурных различий, суть которой заключается в следующем: объекты на первом изображении задаются в виде набора незашумленных уровней яркости; второе изображение строится на основе первого путем добавления к нему новых уровней яркости; далее к обоим изображениям добавляется независимый аддитивный шум с известными параметрами. Такая модель достаточно адекватно отражает возникновение на реальных снимках изменений, связанных с появлением или исчезновением объектов сцены. Также в первой главе показывается связь полученной теоретической модели с морфологическим анализом изображений Ю.П. Пытьева (для случая связей между парами пикселей) и с теорией марковских случайных полей (МСП) [83] (в случае связей групп пикселей).

Заключительная часть главы посвящена математическому исследованию алгоритма, основанного на морфологическом проекторе. В рамках предложенной формализации доказываются утверждения, определяющие распределения вероятности яркости зафиксированного пикселя изображения, представляющего собой разность исходного изображения и его морфологической проекции для случаев, когда в этом пикселе есть, либо нет структурного различия. На основе этих распределений выводится формула оптимального порога алгоритма.

В **третьей главе, “Алгоритмы поиска структурных различий и их реализация”** приводится описание разработанных алгоритмов поиска структурных различий, которые можно разделить на два класса: алгоритмы, основанные на сканирующем окне и функциях преобразования яркости, и алгоритм, основанный на глобальной оптимизации. Применение функций преобразования яркости следует из формулировки задачи: объекты, изменившие цвет, но сохранившие форму, структурными различиями не считаются, поэтому сначала необходимо “выровнять” яркости входных изображений. Это делается с помощью функций преобразования яркости, в качестве которых используются морфологический проектор Ю.П. Пытьева и его модификации, связанные с ограничением допустимого класса преобразований яркости линейными и квадратичными функциями, а также регуляризацией самого проектора. Алгоритмы, основанные на функциях преобразования яркости, состоят из следующих шагов:

1. Исходные изображения сканируются окном заданного размера.

2. Для каждого положения окна производится выравнивание яркостей изображений путем преобразования яркостей фрагмента первого изображения на основе его структуры и яркостей второго изображения.

3. Строится разностное изображение как попиксельная разность преобразованного и второго изображений.

4. Симметризация результата, о сути которой подробно рассказано в пункте 2.1.

5. Пороговая обработка разностного изображения и построение связных областей найденных структурных различий.

Алгоритм, основанный на глобальной оптимизации энергетической функции, для построения выровненных по яркости изображений использует теорию марковских случайных полей (МСП) и оперирует не с локальным окном, а с окрестностью каждого пикселя. Энергетическая функция состоит из двух слагаемых: первое отвечает за близость яркостей двух изображений, второе — за сохранение их структуры. Метод реализован с помощью алгоритма имитации отжига [8]. Также разработан алгоритм, основанный на непосредственном применении введенного определения структурных различий ("степень биективности"). Данный подход полезен для поиска структурных различий малого размера и для устранения локальной рассинхронизации изображений. Для учета возможного локального геометрического несоответствия снимков возможно применение алгоритма с оптимальным выбором положения сканирующего окна на основе вычисления "степени биективности".

Также в данной главе приводятся анализ времени работы алгоритмов, описание программного комплекса, интегрированного в среду визуализации и обработки данных дистанционного зондирования Земли ENVI и внедренного для использования в работу ОАО "НИиП центр Природа". Ввиду вычислительной сложности предложенных методов была разработана их реализация, использующая высокопроизводительные многоядерные графические процессоры [41], в результате которой получено уменьшение времени работы примерно в 6–51 раз (в зависимости от алгоритма).

В **четвертой главе, "Исследование алгоритмов поиска структурных различий"** дано описание вычислительного эксперимента по сравнению результатов работы предложенных алгоритмов. Описывается методика проведения такого эксперимента, в которой используются только реальные снимки с идеальной (освобожденной от влияния оператора) ручной разметкой, и имеется возможность варьировать тип и уровень шума, добавляемого к изображениям, а также выбирать размер и тип объектов, присутствующих на изображениях. Данная методика представляет собой развитие математической модели из первой главы, более близкое к реальным снимкам, но не поддающееся теоретическому исследованию. Помимо сравнения качества работы алгоритмов данная методика позволяет прояснить вопрос выбора размера сканирующего окна алгоритмов. При проведении вычислительного эксперимента был использован суперкомпьютер "Уран" ИММ УрО РАН [45].

В **заключении** подведены основные итоги работы, выносимые на защиту, а также приводятся данные о публикациях и апробациях автора по теме диссертации и обсуждаются некоторые перспективные направления дальнейших исследований.

Глава 1

Задача поиска структурных различий изображений

В данной главе вводятся обозначения, используемые в дальнейшем. Приводится обзор литературы по теме поиска структурных различий на изображениях и указывается место данной работы в ряду предшественников.

1.1 Основные обозначения

Пусть множество пикселей X (поле зрения) — ограниченное подмножество плоскости $\mathbb{R}^2$, $S = |X|$ (количество пикселей) будем называть *размером* (или *площадью*) изображения. Под *изображением* f будем понимать числовую функцию $f : X \to \mathbb{R}$, а значение $f(x)$ будем называть *яркостью* изображения f в точке поля зрения $x \in X$, полагая при этом, что изображения содержат дискретный набор яркостей. Если для некоторой меры μ, заданной на σ-алгебре подмножеств X существует интеграл $\sum\limits_X f^2(x)d\mu < \infty$, тогда *семейством допустимых изображений* будем называть конечномерное евклидово пространство Φ функций f, g, наделенное естественным скалярным произведением $(f, g) = \sum\limits_X f(x)g(x)$ и нормой $\|f\| = (\sum\limits_X f^2(x))^{1/2}$ (более аккуратное определение пространства Φ приведено в [40]).

Также определим класс $\mathfrak{F}$ функций преобразования яркости следующим образом [39]. Для $f \in \Phi$ и $F \in \mathfrak{F}$ выполняются условия: 1) $F \circ f \in \Phi$; 2) если $f(x_1) = f(x_2)$, то $(F \circ f)(x_1) = (F \circ f)(x_2)$ (т. е. F не может разбивать уровни яркости). Здесь $F \circ f$ означает суперпозицию: $(F \circ f)(x) = F(f(x))$.

1.2 Обзор литературы

Задача поиска существенно изменившихся областей (*change detection*) на последовательности изображений вот уже много лет привлекает внимание исследователей по всему миру. Причиной такой популярности является многообразие прикладных задач, для решения которых требуется знание о произошедших в изучаемой сцене изменениях: это задачи видеонаблюдения (в которых поиск различий является предобработкой с целью уменьшить объем входных данных), видеокодирования, анализ движения, медицинские приложения и анализ космоснимков (мониторинг земной поверхности). Вследствие этого разработано и предложено необычайно широкое многообразие практических алгоритмов, позволяющих решать данную задачу. Одна из самых важных проблем при решении задачи поиска различий связана с определением того, что является этим самым различием, т. е. с необходимостью построения специальных математических моделей для существенно изменившихся областей. Поэтому поиск различий более правильно называть поиском *существенных* различий, и каждый алгоритм настраивается на специфику конкретных практических требований.

В общем случае, решение задачи поиска различий состоит из следующих шагов:

1. Геометрическая и радиометрическая коррекция входных изображений.
2. Собственно поиск различий, результатом которого является полутоновое разностное изображение (*difference image*).
3. Пороговая обработка разностного изображения.
4. Оценка точности работы алгоритма.

Каждый из этих пунктов представляет собой отдельную сложную задачу, и точность работы алгоритма поиска различий напрямую зависит от качества выполнения данных операций.

1.2.1 Геометрическая и радиометрическая коррекция входных изображений

По своей сути поиск различий — это работа с последовательностью снимков, полученных в разное время и, следовательно, при различных условиях съемки. Учет и устранение влияния условий регистрации изображений играют ключевую роль в предобработке входных изображений, связанной с разделением существенных и несущественных изменений.

Геометрическая коррекция входных изображений призвана обеспечить их геометрическую выравненность, под которой понимается то, что каждый пиксель x на обоих снимках имеет одни и те же координаты на местности, т. е. эти снимки одинаково привязаны к местности. Если говорить о спутниковой съемке, то геометрическая коррекция выполняется с помощью координатной привязки ключевых точек снимков в разнообразных ГИС с учетом знания параметров спутника, съемки и полученных изображений. Для анализа видеопоследовательностей в подобной коррекции, как правило, нет необходимости из-за фиксированного положения камеры, хотя существуют и исключения. Например, системы помощи водителю [61], где движение и камеры, и объектов делает геометрическую коррекцию невозможной, и вся тяжесть решения задачи поиска различий ложится на основное ядро алгоритма. В целом, если движение камеры незначительно и априорно известно, то геометрическую коррекцию можно свести к некоторому пространственному преобразованию (аффинному, проективному и т. д.) [104].

Отдельные работы посвящены влиянию недостаточной геометрической выравненности (рассинхронизации) изображений на результат поиска различий. В работе [117] приведено описание эксперимента по оценке влияния рассинхронизации для изображения и его копии, благодаря чему исключено влияние структурных различий (аналогичный подход будет использован в главе 4 данной диссертации для сравнения алгоритмов поиска различий), и та же оценка для рассинхронизированных разновременных изображений. Рассинхронизация тождественна сдвигу изображения по обеим осям. Для обоих экспериментов построены графики зависимости величины, представляющей собой относительную сумму квадратов попиксельной разности исходного и сдвинутого изображений (некоторый аналог величины структурного различия, называемый полудисперсией), от величины сдвига. Графики в обоих случаях получились похожими на график степенной функции с показателем 0.5. Результаты таковы: ошибка поиска различий, связанная с рассинхронизацией изображений, составляющая 10% от величины собственно структурных различий, может быть вызвана сдвигом в 0.2 пикселя. При сдвиге в 1 пиксель эта ошибка достигает 50%. Данный подход был развит в работе [58], где авторы выбирали разные интервалы длин волн для эксперимента. Результаты аналогичны статье предшественников [117]: ошибка в регистрации 0.1934 пикселя позволяет найти 90% истинных структурных различий среди всех обнаруженных, а ошибка в 1 пиксель приводит к резкому падению точности.

Необходимость радиометрической коррекции входных изображений обусловлена двумя факторами: 1) поглощением и рассеянием электромагнитного излучения атмосферой Земли; 2) изменением параметров освещенности сцены. Атмосферная коррекция снимков и ее влияние на задачи поиска различий и классификации рассмотрены в работе [112]. Авторы разделяют коррекцию на абсолютную (модель атмосферного рассеяния Релея, учет отражательной спо-

собности объектов и др., прилагается описание четырех методов) или относительную (на основе линейного преобразования яркости по паре входных изображений). Фактический вывод авторов: в задаче поиска различий атмосферная коррекция нужна только при использовании методов, связанных с обучающими выборками, либо при использовании специфических форматов (NDVI [98], например), по причине того, что либо методы используют классификацию изображений как единого целого, либо различия ищутся как попиксельная разность, и тогда влияние атмосферы можно учесть как константу, добавляемую к пороговому значению.

Изменение яркости и/или позиции источника освещения может привести к существенному различию в наборе яркостей соответствующих пикселей разновременных входных снимков [104]. Подобное влияние условий регистрации особенно критично при анализе космоснимков, и для их устранения используются методы относительного выравнивания яркости исходных изображений. В работе [121] методы относительного выравнивания подразделяют на статистические (преобразование с целью сделать похожими значения max/min, среднего, дисперсии), гистограммные (сделать похожими гистограммы) и линейные (линейная функция преобразования яркостей). Общая схема последних состоит в вычислении среднего и дисперсии исходных изображений и нахождении коэффициентов линейного преобразования: наклон есть отношение дисперсий первого снимка ко второму, сдвиг определяется наклоном и математическим ожиданием. Варианты методов связаны с выбором набора пикселей, по которым будут считаться среднее и дисперсия. В методе *Image Regression* этот выбор осуществляется по всем пикселям изображений [99]. В методе *Pseudoinvariant Feature* (PIF) — по пикселям объектов, для которых отражательная способность условно независима от природных условий и яркость достаточно однородна, т. е. это, как правило, антропогенные объекты, которые можно извлечь из инфракрасного диапазона [66,108]. В методе *Radiometric Control Set* (RCS) рассматриваются пиксели из метода PIF (как правило, достаточно яркие) плюс пиксели темные (вода, например) [67]. В методе *No Change* (NC) *Set Determined from Scattergram* на диаграмме рассеяния выбираются точки условных центров воды и поверхности земли без изменений (как локальный максимум на диаграмме рассеяния). По этим двум точкам строится прямая и выбирается окрестность вокруг нее. Для вычисления дисперсии и среднего используются только точки, попавшие в эту окрестность [60]. Сравнение авторами этих методов дало следующие результаты: методы, использующие большое количество пикселей (*Image Regression* и *NC Set Determined from Scattergram*), уменьшают динамический диапазон снимков и приводят к потере важной информации, что критично, если в дальнейшем применять к снимкам, например, текстурную классификацию. Оставшиеся методы диапазон яркостей сохраняют, но сильно затирают собственно структурные различия.

Один из наиболее мощных методов для относительной радиометрической (яркостной, про-

ще говоря) коррекции был предложен в морфологическом анализе изображений Ю.П. Пытьева. В работах [39,40] им были развиты методы, инвариантные относительно преобразований, характеризующих влияние условий регистрации изображений. Центральным понятием морфологического анализа изображений Ю.П. Пытьева является *форма изображения* f, которую можно определить как оператор проектирования P_f изображения g, действующий в пространстве Φ, на множество $V(f) = \{F \circ f \mid F \in \mathfrak{F}\}$. Как показано в [39,40], этот оператор является оператором проектирования и его формула такова:

$$P_f g(x) = \sum_i \frac{\sum\limits_{x' \in X} g(x') \cdot \chi_i^f(x')}{\sum\limits_{x' \in X} \chi_i^f(x')} \cdot \chi_i^f(x), \text{ где } \chi_i^f(x) = \begin{cases} 1, \text{ если } f(x) = i, \\ 0, \text{ иначе,} \end{cases} \tag{1.1}$$

где суммирование по i производится по всем значениям функции f. С содержательной точки, проектор работает следующим образом: он усредняет яркости второго изображения g по уровням яркости первого f.

В работах [5–7] приведено обобщение морфологического подхода в задачах обработки изображений, основанное на построении и оптимизации априорных критериев (функционалов) проектирования, учитывающих специфику решаемой задачи (в отличии от подхода [39,40], опирающегося только на алгебраические свойства оператора проектирования). Формализация понятия структурного возмущения и вопрос устойчивости алгоритмов к нему рассмотрены в [36,102]. В дальнейшем именно морфологический проектор будет использован как основа для предлагаемых в данной работе алгоритмов поиска структурных различий изображений.

1.2.2 Поиск существенных различий

В общем случае, цель задачи поиска различий — это ответ на вопрос: есть ли в данном пикселе значимое различие между двумя изображениями. Соответственно, первой и наиболее распространенной техникой является прямое попиксельное вычитание входных изображений [91,104]. Разумеется, такой подход будет лишен смысла без полноценной геометрической и радиометрической коррекции исходных снимков. При всей своей простоте такой подход является мощным средством поиска практически любых типов изменений в любых прикладных приложениях, но обладает двумя внушительными недостатками: крайняя подверженность влиянию шума и сложность в выборе порогового значения, определяющего существенность найденных изменений. Идейным развитием такого подхода стал метод *change-vector analysis* (CVA), в котором яркость каждого пикселя представляет собой вектор (по сути, это цветные изображения), и, следовательно, яркость пикселей разностного изображения также является вектором. Модуль данного вектора служит индикатором степени различия, а угол наклона можно использовать

для определения типа изменения. Еще один вариант развития подобного подхода предложен в [111] для анализа видеопоследовательностей и называется *shading model*. Суть такова: по области интереса (сканирующему окну) строится выборка как попиксельное отношение яркостей двух фрагментов (что дает устойчивость к освещенности) и для полученного набора чисел рассчитывается среднее и дисперсия. Если дисперсия велика, то утверждается, что обнаружено структурное различие.

Использование видеопоследовательности (что довольно затруднительно при анализе космосъемки земной поверхности) и оценка различий не по всему изображению, а по сканирующему окну описаны в [54]. Здесь для измерения степени различия используется разность исходного изображения и его оценки (для случая устраненного влияния условий регистрации) на основе второго изображения, плюс симметризация (т. е. независимость от порядка изображений). Для построения оценки используются либо линейные функции, либо нелинейные, дающие оптимум в методе наименьших квадратов. Для улучшения качества используется построение временной последовательности изменений. Однако для создания такой последовательности нужно проводить пороговую обработку для разностного изображения каждой пары входных снимков, т. е. нужно выбирать многочисленные пороги. Бинарные изображения после пороговой обработки сглаживаются функцией Гаусса, следовательно, точность локализации различий падает. Тем не менее, поиск различий именно для видеопоследовательностей является крайне популярной задачей, для которой существуют специализированные базы данных и тесты для сравнения качества работы различных алгоритмов [55]. Еще одной важной задачей при работе с видеопоследовательностями является оценка качества их сжатия, для чего используется разработанный для оценки качества сжатия изображений индекс структурного сходства (*structural similarity*, SSIM) [120].

Другой подход к улучшению качества работы простейшей разности снимков — это использование многомасштабных изображений, предложенное в работе [66]. Авторы используют для яркостной коррекции метод PIF по целому снимку. Далее из выровненных по яркости изображений строится набор разномасштабных снимков с помощью алгоритма, вычисляющего дисперсию и среднее, по которым определяются оптимальные размеры объектов и потом формируются границы с помощью разновидности алгоритма “водораздела” (*watershed*) из [69]. Далее вычисляется разность изображений по каждому масштабу и комбинируется итоговый результат.

Тем не менее, в обработке изображений выработан огромный арсенал средств, позволяющих получить из снимков гораздо больше информации, чем просто значения яркостей в пикселях. Речь идет о текстурах, контурах, объектах и т. д. Используя текстуры можно разнообразить

базовые методы поиска различий. В работе [70] авторы предлагают расширить CVA подход за счет использования текстурной информации. Основываясь на работе Харалика [68], авторами предлагаются следующие индексы: контраст, среднее, стандартное отклонение, второй угловой момент, корреляция, однородность, неоднородность и энтропия (последние три в работе не используются). Для пяти оставшихся индексов строятся текстурные изображения, яркости которых представляют собой нормированные значения соответствующих индексов. Далее к этим пяти изображениям применяется *change-vector analisys*. Вместе с яркостным анализом получается двухмерное изображение модулей векторов изменений, анализ и пороговая обработка которого дают конечный результат.

Другой вариант объединения попиксельной разности изображений и текстурной информации описан в [84, 85]. Авторы предлагают в дополнение к измерению отличия по яркости проводить текстурную сегментацию изображений. Для учета текстурной информации используется отношение взаимной- и автокорреляции для соседей пикселя в окрестности размером 5×5. Отличия по яркости — это попиксельная разность исходных изображений. Далее строится общая функция различия как взвешенная сумма яркостных и текстурных отличий. Однако из-за требований к выполнению программы в реальном времени (авторы применяют свой метод для обнаружения движения по видеопоследовательности) глобальную оптимизацию полученной функции не проводят, ограничиваясь некоторым итеративным алгоритмом, минимизирующим ошибку энергетической функции измеренных отличий. В обработке космоснимков данная идея получила свое развитие в работе [75], объединяющей попиксельные и объектные подходы путем сегментации изображений на однородные области с помощью алгоритма *mean shift* [57], которые далее уже рассматриваются как соседи окрестности размером 5×5 пикселей для алгоритма [84, 85]. Разумеется, качество работы алгоритма будет напрямую зависеть от точности сегментации, поэтому авторы предпочитают работать со снимками высокого разрешения. Идея сегментации снимков высокого разрешения для улучшения качества работы алгоритмов поиска различий также использована в работе [82]. Предложенный авторами алгоритм состоит из трех шагов: сегментация; геометрический поиск различий (длина/ширина объекта плюс координаты крайних точек и центра тяжести); текстурный поиск различий (вейвлет преобразование и моделирование коэффициентов с помощью распределения Гаусса, которые сравниваются с помощью метрики Кульбака–Лейблера [81]).

Как уже понятно, набор разнообразных полезных индексов очень и очень широк. В [71] используются энтропия, контраст, матрицы смежности и другие, к которым добавляются индексы теории информации (*mutual information* [64] и количество информации), а выравнивание яркостей изображений делается на основе функции преобразования яркости, аналогичной мор-

фологическому проектору [39, 40] (только берется не среднее, а мода). Далее происходит комбинирование индексов на основе взвешенной суммы, и для принятия решения применяется теория *Dempster–Shafer'a* [59,110]. Для временных последовательностей изображений также можно использовать фильтр Калмана [107]. В работе [50] поиск различий для аэрофотоснимков (как несоответствие строения/растения) осуществляется на основе текстурного анализа и нейронных сетей. Для этого изображения разбиваются на фрагменты 64×64 пикселя, и для каждой пары фрагментов вычисляются текстурные признаки первого порядка: среднее, дисперсия, третий момент, энтропия; второго порядка (матрицы смежности Харалика [68]): корреляция, энтропия, контраст, однородность и т. д.; и более высокого порядка на основе *run-length*-матриц [114]. Разделяющее правило состоит из последовательного выполнения множества базовых алгоритмов (метод ветвей и границ и др.). Для обучения и классификации используется нейронная сеть с 16 нейронами на входе (это признаки различных порядков) и двумя на выходе (строения/растения). Далее фрагменты по 4 штуки объединяются в связные множества. Если во всех из них результаты классификации пары изображений различны, то считается, что структурное различие есть, т. е. авторами решается задача чисто обнаружения различия без его локализации.

Обучающиеся алгоритмы также нашли применение при нахождении различий на разновременных изображениях. Предлагаемый в [118] алгоритм работает с тремя признаками: среднее значение яркости, энергетическая функция текстуры и индекс формы. Все эти признаки извлекают из изображений с помощью пакета *eCognition* [35] (стоит отметить, что на этот пакет ссылаются многие авторы). Каждый из признаков обрабатывается в нескольких масштабах. Выбор ряда масштабов осуществляется на основе предварительного обучения (также с помощью *eCognition*). Для каждого масштаба осуществляется поиск различий, и результаты по признаку суммируются с весом, пропорциональным дисперсии результата. Далее результаты по трем признакам суммируются с весами в отношении 9(яркость):7.5(текстура):3(форма). Другой вариант обучения связан с нечеткими интегралами и представлен в работе [97].

Большая группа методов рассматривает пару входных изображений, содержащих большой количество каналов, как набор случайных векторов, и отыскивает взаимосвязь между ними с помощью статистических методов: анализ главных компонент, методы Грамма-Шмидта, χ-квадрат и другие. Один из наиболее значимых среди них предложен в [95] и основывается на идее канонического корреляционного анализа, преобразуя исходные наборы случайных величин в так называемые канонические случайные величины, корреляция между которыми проявляется более отчетливо. В работе [96] предложена более устойчивая регуляризация данного метода. Использование набора пар изображений для проверки статистических гипотез — существенное различие, несущественное различие и отсутствие различий — описаны в работе [88].

Наконец, под специфику задачи поиска различий также были разработаны особые индексы, например, моменты циклического сдвига (*circular shift moments*) в работе [89]. Данные моменты вычисляются по областям интереса (*area of interest*, AOI, фактически сканирующее окно, только малого размера 5×5 пикселей) и делятся по направлениям: по x и по y. Например, момент по x — это отношение взвешенной суммы яркостей столбцов AOI к общей сумме яркостей по AOI. Вес столбцов нулевого порядка равен просто номеру столбца (считая слева направо). Для момента первого порядка первый столбец получает вес N (это общее число столбцов), второй столбец получает вес 1 и т. д. По y аналогично. Поэтому моменты названы циклическими, хотя в статье авторы используют только моменты нулевого порядка и приводят итеративную формулу вычисления этих моментов, начиная с нулевого. Выбранные моменты вычисляются для обоих сравниваемых изображений, после чего берется их разность. В случае отсутствия на изображениях и шума, и структурных различий такой подход дает нулевую разность (или близкую к нулю). В случае присутствия шума находится AOI, дающая минимум суммы модулей разностей моментов по x и по y — это область, где наиболее вероятно отсутствие различий. По этой AOI обоих изображений с помощью специальной формулы вычисляется дисперсия шума (модель шума стандартная: аддитивный, независимый, Гауссов с математическим ожиданием ноль). Далее вычисляется дисперсия момента для каждой области как отношение дисперсии шума к квадрату суммы яркостей пикселей AOI, умноженной на коэффициент, зависящий от размеров AOI. Сумма таких дисперсий по фрагментам обоих изображений дает оценку дисперсии разности моментов. Выдвигаются две гипотезы: нулевая, что математические ожидания этих моментов равны (различий нет), и первая, что они не равны и различия есть. Задается уровень значимости, и если модуль разности моментов по x или по y превзошел отношение стандартного отклонения разности моментов к этому уровню, то структурные различия есть. Метод получился сложным, но по представленным результатам неплохо работает для сильно зашумленных видеопоследовательностей.

Поиск структурных различий на основе анализа контуров является мощным методом, позволяющим успешно отсеивать различия, связанные со сменой условий регистрации, изменением угла съемки и т.д. Для поиска непосредственно контуров можно использовать либо подход из [53], либо искать не прямые, а сразу замкнутые контуры [29] с помощью преобразования Хафа [72]. В работе [86] предлагается алгоритм, сравнивающий не яркости областей, а их границы, что позволяет автоматически устранять влияние условий регистрации. Алгоритм состоит из следующих шагов:

1. Регистрация изображений. Предполагается, что дороги, площади, поверхность земли в случае отсутствия на них изменений остается на месте на обоих снимках. Но могут сдвинуться

точки, находящиеся выше — крыши домов. Для таких точек вычисляется некоторое Главное направление.

2. Поиск прямых линий с помощью метода [53].

3. Из п. 2 находятся прямые, соответствующие различиям, путем сравнения их перпендикуляров (отклонение по порогу) с *multi-scale gradient orientation field* (GOF — это изображение, каждая точка которого получена путем свертки ориентации градиента с функцией Гаусса).

4. Для устранения различия в углах съемки производится сопоставление ключевых точек алгоритма SIFT [90] по максимуму корреляции, малому отклонению для точек на поверхности земли и для "высоких" точек — по совпадению направления с Главным направлением. Далее, используя масштабный фактор метода SIFT, отсекаются лишние линии.

5. С помощью математической морфологии ищутся тени и удаляются линии либо лежащие на границах тени, либо лежащие внутри теней на втором снимке.

6. Группировка линий и отсечение малых по площади областей. Два отрезка попадают в кластер, если расстояние между любой парой их концов меньше заданного порога.

Фактически, в статье [86] описаны все сложности, встречающиеся в задаче поиска различий: изменение условий регистрации, структурный шум (машины и т. д.), изменение угла обзора и наличие теней. Радиометрическая коррекция снимков из предыдущего пункта позволяет уменьшить влияние условий регистрации. Структурный шум, т. е. несущественные для данной задачи изменения, можно отсеивать по размеру, но это добавляет новую проблему — выбор подходящего порога. Таким образом, этот вопрос оказывается на стыке определения модели, требований к снимкам и алгоритмических решений. Изменение угла обзора исправить возможно лишь полной и очень качественной сегментацией, что также является крайне сложной задачей. Наконец, исключение теней также требует отдельных алгоритмов (например, [115]). Все вышеперечисленное, вместе с требованиями конкретных приложений в различных сферах применения, делает чрезвычайно затруднительным создание хотя бы относительно универсального алгоритма для поиска различий, поэтому многие авторы в своих работах предлагают некоторые специфические модификации алгоритмов.

Возможно, наиболее очевидный вариант — это интеграция алгоритма поиска различий с информацией ГИС [44] (различными картами и планами). В работе [49] автор, опираясь на алгоритм [52], осуществляет многомасштабную сегментацию изображений и строит иерархическую структуру связей между объектами. Далее результаты сегментации улучшаются путем добавления информации, полученной из различных ГИС и пакета *eCognition* [35]. Сравнение карты местности и космоснимка также применяется в [78]. Здесь карта (хотя авторы утверждают, что также можно использовать результаты сегментации изображений) задает границы

областей, внутри которых по пикселям снимка считаются среднее и дисперсия. Далее вычисляется отклонение яркости каждого пикселя от среднего (относительно дисперсии), и все пиксели разбиваются на четыре группы по степени вероятности присутствия различий. Похожий подход использован в [105], где поиск различий производится на основе размеченной карты высокого разрешения (за счет чего достигается субпиксельная точность) и серии снимков низкого разрешения с необычным подходом к построению модели на основе отсутствия структуры.

Если постановка задачи сводится к экологическому аспекту анализа растительного покрова, то наиболее эффективным становится применение NDVI-серий снимков высокого разрешения [98]. В работе [122] для таких снимков сначала используется стандартный *change-vector analisys*, результаты которого для исключения изменений, связанных с модификациями типа изменения вида растительности, обрабатываются с помощью нормированной кросс-корреляционной функции [119]. В последней работе нормированная кросс-корреляционная функция считается для графиков зависимости индекса NDVI от времени с некоторым параметром сдвига. Выбирается значение сдвига, дающего максимум корреляции, после чего по набору обучающих примеров применяется пороговая обработка. Показано, что корреляция почти не подвержена влиянию возможных загораживаний (облака) и шума.

Оригинальный подход к поиску различий для аэрофотоснимков состоит в сравнении последнего по времени изображения (все параметры съемки которого считаются известными) не с более ранним снимком, а с $3D$-моделью сцены, в которой здания составлены из прямоугольных параллелепипедов [76]. Сначала авторы извлекают отрезки прямых из снимка и проводят их сопоставление с моделью. Для этого делается предположение о том, что модель и снимок отличаются только на сдвиг, и с помощью известных параметров съемки производится сопоставление. После этого происходит подтверждение модели. Далее вычисляется значение функционала, представляющего собой взвешенную сумму следующих параметров (каждый от 0 до 1): видимость линии (из модели и известного положения камеры), наличие линий (т. е. пересечение линий изображения с краями на модели), покрытие линий (на сколько полно линии пересекаются), наличие углов (отношение количества найденных углов изображения к их количеству, рассчитанному по модели), наличие теней (отношение количества углов и границ теней изображения к их количеству по модели; тени находятся с помощью метода [87]). Если полученное значение велико (больше 0.5), то различий нет; если меньше 0.2, то различия есть; среднее значение указывает на необходимость дополнительной информации (фактически нового снимка); после этого производится обновление модели. В целом, авторы очень точно наметили все трудности, возникающие при таком подходе к поиску различий, и утверждают, что их метод с ними успешно справляется.

Системы анализа движений применяются в различных приложениях, и одним из актуальнейших в настоящее время является система помощи водителю. Поиск структурных различий для такой системы основан на применении нейронных сетей и разделов психологии, связанных с познавательными процессами сознания человека. Метод, предложенный в [61], работает с цветной видеопоследовательностью. Поскольку движутся все попадающие в кадр объекты, то авторы фокусируются на поиске дорожной разметки, туннелей и мостов. Собственно поиск различий осуществляется следующим образом: по каждому каналу видеопоследовательности строятся два изображения, соответствующие максимальной и минимальной яркости каждого пикселя. Разностное изображение строится как разность этих двух изображений. Далее строится дифференциальное изображение как разность последовательных разностных изображений. Каждое дифференциальное изображение подается на вход первой нейронной сети, отвечающей за фокусировку внимания и извлечение ключевых признаков (яркость, цвет, текстуры и т. д.; для этого как маски используются разностные изображения). Далее данные передаются во вторую нейронную сеть, где происходит распознавание обнаруженного объекта и обучение (с учителем, если признаки совпадают с заранее определенными, или без учителя в противном случае). Результат работы алгоритма — выделенное изменение и его тип.

Поиск различий для снимков сенсорами радаров с синтезированной апертурой представлен в [56]. Для этого авторами моделируется многомерное Гамма-распределение, и оценка его коэффициентов корреляции лежит в основе поиска различий. О распараллеливании алгоритма поиска различий написано в [103]. Здесь авторы работают исключительно со снимками MODIS [43] (сенсора на спутниках Терра и Аква), поскольку считывают несколько длин волн. Сам алгоритм поиска различий представляет собой разность изображений по каждой длине волны. Наконец, в работах [91, 104] представлены классификации еще нескольких десятков методов, разработанных для поиска различий на разновременных снимках.

1.2.3 Пороговая обработка и оценка точности работы

Результатом применения выбранного метода поиска различий является полутоновое *разностное изображение* (*difference image*), в котором яркость характеризует уровень (или степень) найденного в данном пикселе изменения. Последний этап алгоритма нахождения различий — это построение на основе разностного изображения связных областей, соответствующих существенным изменениям. Простейшим путем получения этих областей является пороговая обработка, однако, в условиях зашумленности входных, и, как следствие, разностного изображений, выбор подходящего порогового значения становится крайне нетривиальной задачей. Как правило, этот порог выбирается вручную на основе опыта и знаний оператора, что позволяет достигнуть хо-

роших итоговых результатов. В то же время в литературе разработан широкий спектр методов автоматического выбора порогового значения.

По своей сути, пороговая обработка разностного изображения — это задача классификации пикселей на две группы: содержащие существенные различия и не содержащие таковых. Следовательно, вполне оправданным становится применение классических алгоритмов: *support vector machine* в статье [70] (для его обучения авторами использовалось 30 положительных и 25 отрицательных примеров); метод главных компонент [63] (с привлечением теории нечетких множеств для комбинирования результатов поиска по разным каналам); расстояние Махаланобиса в роли разделяющей функции [58].

Другой путь предложен в работе [92] и связан с выбором некоторых критериев, в качестве которых авторы предлагают три варианта: *Kittler and Illingworth's algorithm* (основанный на Байесовской теории) [77], *Huang and Wang's algorithm* (теория нечетких множеств) [74] и *Otsu's algorithm* (в предположении о нормальности распределений для классов различие/нет различия) [101]. Второй класс критериев основан на алгоритме *Expectation–Maximization* [94]. Приводится сравнение алгоритмов: из первого класса лучшие результаты показал алгоритм *Kittler and Illingworth's*, затем алгоритм *Otsu* (точность, примерно, 95%). Метод *Huang and Wang's* дает почти 50% ложных тревог. Во второй группе варианты алгоритмов имеют точность, примерно равную 97%.

Моделирование распределений шума и яркостей исходных изображений также является широко распространенным подходом к задаче выбора оптимального значения порога. В работе [106] выделяется четыре принципа выбора порога:

1. На основе модели яркости шума (распределение Гаусса, дисперсию предлагается оценивать с помощью метода наименьших квадратов).

2. На основе модели пространственного распределения шума (распределение Пуассона).

3. Распределение яркости исходного сигнала (аппроксимация распределения яркости для окна либо заранее определенной функцией [73], либо использованием непараметрических методов типа критерия Колмогорова–Смирнова).

4. Пространственное распределение исходного сигнала (построение графика зависимости устойчивого числа (характеристики) Эйлера от порога; оптимальной будет величина в точке перегиба графика).

В результате сравнения наиболее многообещающие результаты были показаны пространственными методами (распределение Пуассона и числа Эйлера).

Метод, объединяющий черты вышеперечисленных подходов, предложен в [51]. Здесь разностное изображение размечается по области интереса AOI (сканирующему окну) метками

(есть/нет структурное различие), и производится минимизация специальной энергетической функции, представляющей собой сложение по всем пикселям AOI суммы двух функций: первая отвечает за контекст, вторая — за данные. *Контекст* — это взаимодействие метки текущего пикселя с метками соседей, основанный на определении марковских случайных полей [83] и функции Кронекера. *Данные* — взаимодействие яркости пикселя с его меткой, построенное путем моделирования распределения яркости разностного изображения на основе его гистограммы. Моделирование включает два шага: сначала непараметрическая оценка с помощью алгоритма *Reduced Parzen estimate*, взятого из [62], и затем итеративная процедура улучшения оценки с помощью алгоритма *Expectation–Maximization* [94] (в ранних работах авторы предлагали параметрическое моделирование на основе гауссоподобных функций). Оптимизацию итоговой энергетической функции для скорости проводят итеративным алгоритмом *Besag's iterated conditional modes* [48], но можно использовать и метод имитации отжига. Предложенный метод демонстрирует лучшие результаты, чем подобранный вручную порог.

Выбор способа определения оптимального порога алгоритма также является интересной задачей, сложность которой определяется необходимостью подбора критерия точности, т. е. сравнения полученных областей существенных различий с ручной разметкой. В работе [104] предлагается несколько вариантов критерия оптимальности, каждый из которых представляет собой некоторую комбинацию вероятностей ошибок первого и второго рода: *percentage correct classification*, *Jaccard coefficient* и *Yule coefficient*.

1.3 Существующие подходы к определению структуры

Существует множество различных определений структуры изображений, среди которых можно выделить следующие три: морфологическая теория Серра [109], теория многомасштабного анализа [79] и теория морфологического анализа Ю. П. Пытьева [39, 40].

Теория математической морфологии была разработана французскими исследователями Ж. Матероном и Дж. Серра в 1960–1980 гг. В основе математической морфологии лежат теоретико-множественный формализм и представления интегральной геометрии. Ключевым объектом анализа (т. е. носителем структуры) является бинарное изображение, а основным предметом изучения являются морфологические фильтры: эрозия, дилатация, открытие и замыкание. Эти операции используются для устранения шумов на изображениях: с помощью операций эрозии и открытия отфильтровываются "хвосты" объектов нужного размера, а с помощью операции замыкания заливаются внутренние "дырки". Можно сказать, что с точки зрения морфологии два изображения имеют одинаковую структуру, если результат применения некоторой

фиксированной морфологической операции (например, открытия или замыкания с кругом фиксированного радиуса) к этим множествам совпадает. Для сравнения двух структур используется функция расстояния Хаусдорфа.

Теория многомасштабного анализа (Scale-space theory) возникла в 1980-х годах в компьютерном зрении и цифровой обработке сигналов как формальная теория для анализа изображений в различных масштабах. Мотивация создания теории масштабного представления данных основывается на наблюдениях того, что основные объекты реального мира состоят из различных структур разного масштаба. Это означает, что объекты реального мира, в отличие от идеализированных математических объектов, таких как точки или линии, могут проявиться по-разному в зависимости от масштаба наблюдения. В компьютерном зрении нет способа оптимального выбора масштаба для описания структур изображения. Следовательно, единственный разумный подход заключается в рассмотрении описаний в различных масштабах для того, чтобы включить все необходимые вариации масштаба. В основе теории лежит представление изображения в виде однопараметрического семейства изображений, сглаженных с некоторым ядром (обычно это семейство гауссовых функций). *Структурой изображения* с точки зрения многомасштабного анализа является множество светлых и темных пятен в различных масштабах. Функция сравнения двух структур вычисляется при комбинаторном сопоставлении наборов точек из разномасштабных пятен, учитывающем разницу площадей и неточности локализации их центров.

Морфологический анализ изображений появился в конце 1970-х годов в работах Ю. П. Пытьева. В его основе лежит понятие *формы изображения* f, под которой понимается оператор проектирования (морфологический проектор, формула (1.1)) P_f в пространстве изображений Φ на множество изображений, полученных из исходного путем применения всех возможных преобразований яркости: $V(f) = \{F \circ f \mid F \in \mathfrak{F}\}$ [39]. Задав на множестве форм изображений частичный порядок, можно разделить все пространство Φ на классы изображений с эквивалентными формами. Поэтому понятие формы можно интерпретировать как структуру изображения. При этом, вводя разные классы преобразований яркости, можно получать разные типы структур. В этом случае два изображения имеют одинаковую форму (структуру), если одно может быть получено из другого с помощью некоторого преобразования яркости. В работах [5, 7] приведено обобщение морфологического подхода в задачах обработки изображений, основанное на построении и оптимизации априорных критериев (функционалов) проектирования, учитывающих специфику решаемой задачи (в отличии от подхода [39, 40], опирающегося только на алгебраические свойства оператора проектирования).

Помимо вышеперечисленных теорий, структуру изображения можно определить как набор линий контуров, либо как набор множеств пикселей, каждому из которых поставлено в

соответствие значение некоторого дескриптора: это может быть результат текстурной классификации, статистические характеристики и т. д. [11]. Как было показано выше, каждый из этих методов в определенной ситуации используется для решения задачи поиска различий. Таким образом, определение структуры изображения напрямую зависит от постановки задачи и типа искомых изменений.

1.4 Место данной работы в сложившейся системе методов

В вышеперечисленной системе методов решения задачи поиска структурных различий изображений место данной работы представляется следующим образом. Исследование не затрагивает выскоуровневые методы поиска — текстурные, контурные и другие, а рассматривает класс алгоритмов, основанных на предварительном относительном яркостном выравнивании исходных изображений для устранения изменений, связанных со сменой условий их регистрации или наличием несущественных различий, и их последующего попиксельного вычитания для обнаружения изменившихся областей. Этот класс включает в себя следующие подзадачи:

1. Геометрическая коррекция, как правило, считается выполненной. Однако, также предложен вариант алгоритма ее проведения на основе специального критерия, названного “степенью биективности”. При этом методы сглаживания изображения (будь то простое усреднение или более мощные подходы, например, фильтр Винера) в данной работе не рассматриваются (кроме небольшого замечания в п.3.2.1), полагая, что все возможные процедуры по улучшению качества входных снимков проведены заранее.

2. Радиометрическая коррекция представляет собой локальное преобразование в пределах сканирующего окна, которое имеет смысл «относительного яркостного выравнивания» входных изображений. Такое относительное яркостное выравнивание по существу аналогично морфологическому проектированию [40] одного изображения на форму другого. Для этого используются морфологический проектор Ю. П. Пытьева и линейные функции преобразования яркости. На их основе предложены и проанализированы регуляризованный вариант морфологического проектора и квадратичные функции преобразования яркости. Также предложена модификация морфологического проектора, работающая с многоканальными изображениями. Помимо этого, предлагается оригинальный вариант устранения влияния условий регистрации снимков, основанный на теории марковских случайных полей, а также обобщение всех вышеперечисленных подходов.

3. Поиск различий осуществляется путем прямого попиксельного вычитания скорректированных по сканирующему окну изображений.

4. Пороговая обработка исследуется по двум направлениям: во-первых, теоретическое исследование морфологического проектора, основанное на моделировании яркостей исходных изображений и добавляемого шума, и, во-вторых, эмпирическое исследование оптимальных параметров алгоритмов.

5. Поскольку предложенные в [104] критерии точности работы алгоритма показали практически одинаковые результаты, то в качестве итогового был выбран критерий *percentage correct classification*, т. е. минимум суммы ошибок первого и второго рода.

В данной работе основное внимание уделяется подзадачам радиометрической коррекции входных изображений, включая разработку и исследование алгоритмов, и обоснованием выбора оптимального порогового значения. С точки зрения существующих обзоров алгоритмов поиска структурных различий, методы, приведенные в диссертации, относятся к классу алгебраических из [91], и повторяют общую схему решения из [104] с простым вычитанием как методом непосредственного обнаружения изменений.

Также, приводимое ниже определение структуры изображения будет опираться на морфологический анализ изображений Ю. П. Пытьева [39, 40].

Глава 2

Модель структурных различий и задача их локализации

В данной главе приводятся основные теоретические результаты диссертации, преимущественно опирающиеся на теорию морфологического анализа изображений Ю. П. Пытьева [40]. Здесь для корректного решения задачи поиска различий будут сформулированы определения структуры изображений и структурных различий — близкие по сути к морфологии Ю. П. Пытьева, определена математическая модель их описания и поставлена задача их нахождения. После этого будет приведено краткое описание конкретного алгоритма обнаружения структурных различий, основанного на применении морфологического проектора (1.1), для которого будет представлено строгое решение поставленной задачи.

В первом пункте будет дано определение структуры изображения, очевидно, долженствующее лежать в основе всех дальнейших рассуждений. *Структура изображения* — это некоторый способ описания информации о представленной на изображении сцены. Разумеется, для данного снимка таких описаний может быть бесчисленное множество, и выбор конкретного описания напрямую зависит от требований задачи. Структура изображения должна учитывать существенные изменения сцены и быть инвариантной к несущественным в контексте данной задачи изменениям. Помимо этого должен существовать способ сравнения структур изображений между собой с целью определения наличия *структурных различий* между ними.

2.1 Определение структуры изображения

Как уже говорилось ранее, конкретные модели структур и структурных различий правильнее всего привязывать к требованиям поставленной практической задачи, поскольку уни-

версального определения структуры на все случаи жизни дать невозможно. Само по себе понятие структуры не является строго формализованным математическим понятием, а относится к разделу философии, где ему дается следующее определение: структура (от лат. structura — “строение”) — это внутреннее устройство чего-либо, связанное с категориями целого и его частей путем установления совокупности связей между частями целого объекта. Выявление этих связей, изучение взаимодействия и соподчиненности составных частей различных по своей природе объектов позволяет выявить аналогии в их организации и изучать структуры абстрактно без связи с реальными объектами.

На основе этого чисто философского понятия путем уточнения деталей можно сконструировать подходящее математическое определение, пригодное для решения задачи поиска структурных различий изображений. Применительно к анализу изображений категория целого, очевидно, соответствует всему изображению, а его частями являются отдельные пиксели, границы, области однородной яркости, объекты и т. д. Каждый из этих элементов нашел свое применение в алгоритмах, решающих задачу поиска различий на изображениях. В данной работе будет рассмотрен случай, когда формирование структуры будет обусловлено взаимодействием пикселей изображения. Поскольку один пиксель не образует никакой связи, то вполне логично рассматривать неупорядоченные пары пикселей, однако, в общем случае взаимодействовать могут и группы пикселей, вследствие чего рассуждения будут относиться к различным окрестностям пикселя (четырех- или восьмисвязным, например).

Определение 2.1. Под *структурой* Ω изображения f будем понимать семейство множеств $\{L_f\}$, $L_f \subset X$, элементами которых являются взаимодействующие между собой пиксели: $\Omega(f) = \{L_f \mid \forall x_1, x_2 \in L_f \;\; x_1 \leftrightarrow x_2\}$.

Выбор пикселей в качестве носителей структуры обусловлен двумя факторами. Во-первых, предлагаемые в третьей главе данной диссертации методы относятся к группе попиксельных алгоритмов (если рассматривать их относительно классов методов из обзора литературы). Во-вторых, такой подход более гибок и не требует отдельного определения понятий “линия” или “объект”, но в то же время позволяет перейти к ним в случае необходимости. Переход этот устанавливается с помощью конкретизации понятия *взаимодействия* пикселей.

Содержательная суть определения 2.1, задающая ядро структуры, состоит в выборе способа взаимодействия пикселей, который должен отражать основные требования к задаче. В общем случае, взаимодействие между пикселями сводится к двум аспектам: геометрическому и яркостному. Первый отвечает за взаимное расположение пикселей на координатной сетке, второй — за близость значений яркостей пикселей. Соответственно, возможны следующие варианты их

комбинирования:

1. Геометрическая связь, характеризующаяся пространственной близостью пикселей, яркости которых существенно различны. Собственно, это присутствующие на изображении границы:

$$\Omega_{\text{геом}}(f) = \{L_f \mid \forall x_1, x_2 \in L_f \;\; x_1 \sim x_2, \;\; f(x_1) \nsim f(x_2)\}.$$

2. Яркостная связь, состоящая в близости значений яркости пикселей независимо от их пространственного расположения. Такая связь соответствует множеству уровней яркости изображения:

$$\Omega_{\text{ярк}}(f) = \{L_f \mid \forall x_1, x_2 \in L_f \;\; f(x_1) \sim f(x_2)\}.$$

3. Отсутствие связи, т. е. пиксели, далекие друг от друга как пространственно, так и по яркости:

$$\Omega_0(f) = \{L_f \mid \forall x_1, x_2 \in L_f \;\; x_1 \nsim x_2, \;\; f(x_1) \nsim f(x_2)\}.$$

В общем случае эти комбинации дают различное устройство структуры изображения, причем эти построения будут неприводимы друг к другу. Как уже говорилось ранее, выбор конкретного определения должен быть связан с практическими требованиями задачи. В данной работе поиск структурных различий осуществляется с целью оценки площади произошедших изменений, следовательно, наиболее подходящим является определение с яркостной связью между пикселями, т. е. носителем структуры будут множества уровней яркости изображения. Задав уровень яркостной близости $f(x_1) \sim f(x_2)$ как равенство яркостей $f(x_1) = f(x_2)$, получим итоговое определение:

Определение 2.2. *Структурой* Ω изображения f называется семейство $\pounds_f = \{L_f(i)\}_i$ множеств уровня $L_f(i) = \{x \in X \mid f(x) = i\}$ функции f.

Имея определение структуры изображения, можно перейти к различию в этих структурах, т. е. к понятию структурного различия. Кажется очевидным, что для выбранного теоретико-множественного подхода различие в структурах изображений f и g можно задать как симметрическую разность $\{\Omega(f) \triangle \Omega(g)\}$. Однако, такое определение приводит к неоднозначности трактовки структурного различия. Пусть для точек $x_1, x_2 \in X$ выполнено следующее: на первом изображении пиксели принадлежат одному уровню яркости, т. е. $x_1, x_2 \in L_f(i)$, а на втором — разным: $x_1 \in L_g(i_1)$ и $x_2 \in L_g(i_2)$. Это означает, что уровень яркости $L_f(i)$ на втором изображении оказался разбит на два уровня. Следуя установленному пониманию существенного изменения между изображениями, данная ситуация трактуется как появление структурного различия. Но трактуется неоднозначно, поскольку возможны четыре варианта определения структурного различия:

1. Структурное различие соответствует точкам уровня яркости $L_g(i_1)$ (рис. 2.1a).

2. Структурное различие соответствует точкам уровня яркости $L_g(i_2)$ (рис. 2.1b).

3. Структурное различие соответствует точкам обоих уровней яркости $L_g(i_1) \bigcup L_g(i_2)$ (рис. 2.1c).

4. Структурное различие отсутствует.

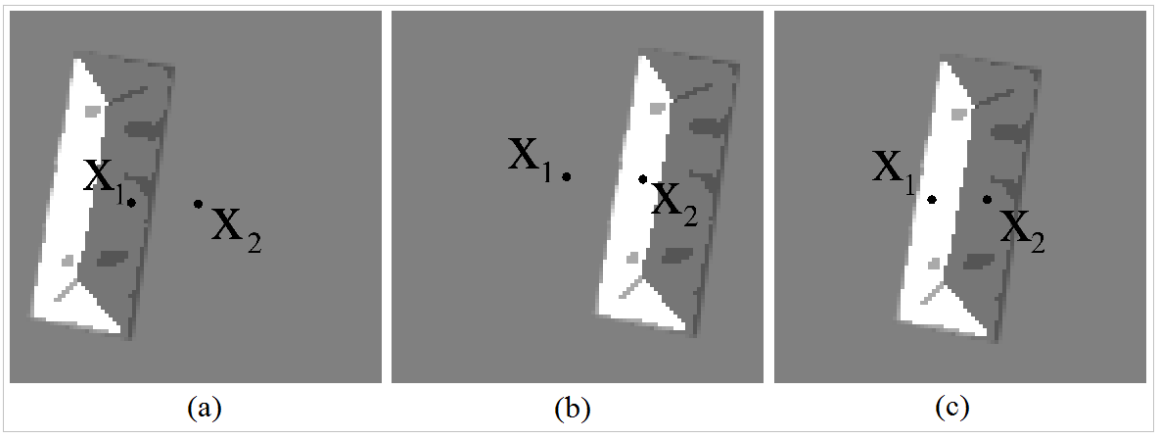

Рис. 2.1. Пример неоднозначного определения структурного различия. Первое изображение считается однотонным, на втором изображении появляется объект, формирующий структурное различие: (a) — в точке x_1, (b) — в точке x_2, (c) — в точках x_1 и x_2 одновременно.

Примеры реализации ситуаций из пунктов 1–3 приведены на рисунке 2.1. Каждый из них связан с появлением нового объекта на втором изображении, но различие состоит в локализации пикселей x_1 и x_2, приводящее к противоречивым вариантам определения структурного различия: либо в одном из пикселей (рис. 2.1a и 2.1b), либо в обоих сразу (рис. 2.1c). Пункт 4 характеризует более сложную и интересную для прикладных приложений ситуацию, связанную с тем, что на реальных снимках в обязательном порядке присутствует шум, искажающий исходные уровни яркости изображений. Исследованию данного случая посвящен пункт 2.6 данной главы.

Пример, изображенный на рис. 2.1, сигнализирует о необходимости использования дополнительной информации, без которой точное определение структурного различия только на основе теоретико-множественных характеристик понятия структуры изображения невозможно. Вполне естественным является шаг, состоящий в присвоении каждому уровню яркости L_f некоторой метки $l : L_f \to \{L_f, l\}$. Тогда наличие структурного различия определяется несовпадением меток для соответствующих пикселей различных изображений. Выбор “пространства” меток связан с выбором алгоритма поиска структурных различий. Например, если алгоритм опирается на текстурную классификацию, то в качестве меток будут использоваться иденти-

фикаторы классов объектов: дом, лес, дорога и т. д. В данной работе исследуются алгоритмы, опирающиеся на методы, работающие исключительно с яркостями отдельных пикселей, и, соответственно, метки l будут выбираться из допустимого диапазона яркостей.

При таком подходе возникает сложность, связанная с объектами, поменявшими свой цвет, а также с изменениями в условиях освещенности при съемке. Ранее упоминалось, что для устранения влияния подобных факторов применяется относительная радиометрическая (или, проще, яркостная) коррекция исходных снимков. В следующем пункте приводится описание общего подхода к яркостному выравниванию изображений и основанная на нем окончательная формализация понятия структурного различия.

2.2 Обобщенная яркостная коррекция изображений

В обзоре литературы упоминались две функции, пригодные для яркостной коррекции изображений: линейная функция и морфологический проектор. Каждая из них связана с идеей сохранения структуры одного из изображений при изменении яркости его пикселей максимально близко к яркостям второго изображения. В данной работе к ним будут добавлены еще две функции, опирающиеся на тот же подход: квадратичная функция и кумулятивная функция особого вида, связанная с учетом влияния соседей каждого пикселя.

Для учета эквивалентности структур двух изображений введем стандартные характеристические функции, которые будут отражать тот факт, что изображение h получено путем изменения значений уровней яркости изображения f с сохранением его структуры. В работе будут рассматриваться алгоритмы, в основе которых лежат следующие характеристические функции структурной эквивалентности:

1. Характеристическая функция $\omega_\pi : \omega_\pi(h; f) = 0$, если существует некоторая функция преобразования яркости $F : \mathbb{R} \to \mathbb{R}$ такая, что $h(x) = F(f(x))\ \forall x \in X$.

2. Характеристическая функция $\omega_1 : \omega_1(h; f) = 0$, если $h(x) = kf(x) + b\ \forall x \in X$ и для некоторых вещественных k, b, и иначе равна бесконечности.

3. Характеристическая функция $\omega_2 : \omega_2(h; f) = 0$, если $h(x) = kf(x)^2 + bf(x) + c\ \forall x \in X$ и для некоторых вещественных k, b, c, и иначе равна бесконечности.

4. Характеристическая функция ω_m:

$$\omega_m(h; f) = \lambda \sum_{x \in X} \sum_{y \in N_x} G_f(h(x) - h(y), f(x) - f(y)),$$

$$G_f(i, j) = \begin{cases} \min(i^2, \alpha^2), & \text{если } |j| < \alpha, \\ 0, & \text{иначе,} \end{cases}$$

где N_x — окрестность точки x, λ и α — положительные параметры. Параметр λ отвечает за близость приведения изображения f к g: чем он больше, тем больше алгоритм будет стремиться сохранить структуру изображения f, из–за чего яркостное приведение f к g будет менее точным. Параметр α характеризует структуру изображения f, задавая допустимый перепад яркости между различными линиями уровня: если яркости двух точек отличаются меньше, чем на α, то они принадлежат одному уровню.

Функции ω_π, ω_1 и ω_2 соответствуют алгоритмам поиска различий на основе поточечной функции преобразования яркости, а функция ω_m соответствует алгоритму поиска различий на основе глобальной минимизации, в реализации которого используется теория марковских случайных полей.

В применении к задаче поиска структурных различий целесообразно использовать характеристические функции структуры для построения изображений f' и g', которые были бы по структуре близки к f и g, а по яркости f' близко к g, а g' близко к f:

$$\begin{aligned} f' &= \arg\min(\|h-g\|^2 + \omega(h;f) \mid h \in \Phi), \\ g' &= \arg\min(\|h-f\|^2 + \omega(h;g) \mid h \in \Phi). \end{aligned} \tag{2.1}$$

После этого вычисление различий осуществляется как нахождение разности построенных изображений с g и f соответственно. Формулы (2.1) будем называть *обобщенной яркостной коррекцией изображений.*

Необходимость построения второго изображения связана с тем, что операция яркостной коррекции несимметрична, т. е. ее применение позволяет найти лишь “появившиеся” изменения (объекты, отсутствующие на первом изображении, но присутствующие на втором). Нахождение “исчезнувших” объектов требует проведения симметризации яркостной коррекции. Теперь можно сформулировать окончательное определение структурного различия:

Определение 2.3. *Структурным различием* называется множество пикселей, удовлетворяющее следующему условию:

$$SD(f,g) = \{x \in X \mid |f'(x) - g(x)| > T \text{ или } |f(x) - g'(x)| > T\},$$

где f' и g' определены формулами (2.1), T — некоторое пороговое значение.

В случае идеальных (незашумленных) изображений $T = 0$, а для зашумленных изображений выбор порогового значения становится отдельной задачей, подходам к решению которой посвящены п.2.6 и четвертая главы диссертации.

Мерой структурного различия изображений будет мощность этого множества. Стоит отметить, что Д. С. Переваловым в работах [36,102] введено понятие *структурной функции* $s(f,g)$ как меры структурного различия изображений, обладающей следующими свойствами:

1. $s(f,g) \geq 0$.

2. Изображения f и g имеют одинаковую структуру тогда и только тогда, когда $s(f,g) = s(g,f) = 0$.

3. Нули функции $s(f,g)$ обладают свойством транзитивности: если $s(f,g) = s(g,r) = 0$, то $s(f,r) = 0$.

Например, для определения 2.3 структурной функцией является $s(f,g) = \max(\|f' - g\|, \|g' - f\|)$, где f' и g' определяются по формулам (2.1). Для морфологического проектора структурную функцию можно определить как расстояние в пространстве изображений между первым изображением f и формой второго $V(g)$: $s(f,g) = \mathrm{dist}(f, V(g))$.

2.3 Структура изображения и морфологический проектор

Определение структуры изображения 2.2 наиболее близко по сути к форме изображения в морфологическом анализе Ю. П. Пытьева. Ключевое отличие между ними состоит в том, что определение 2.2 опирается на теоретико-множественный подход, необходимый для формализации структурных различий, а морфологический анализ сводится к функциональному подходу, важному для доказательства ключевых свойств морфологического проектора. В данном пункте устанавливается связь между введенным понятием структуры изображения, удобным для теоретических исследований, и оператором морфологического проектирования, позволяющим строить прикладные алгоритмы. Следующая теорема показывает, что применение формул обобщенной яркостной коррекции (2.1) эквивалентно построению морфологической проекции Ю. П. Пытьева.

Теорема 2.1. 1. *Метод построения морфологической проекции эквивалентен задаче минимизации*

$$f' = \arg\min(\|h - g\|^2 + \omega_\pi(h; f) \mid h \in \Phi),$$

где $\omega_\pi(h; f) = 0$, *если существует функция преобразования яркости* $F : \mathbb{R} \to \mathbb{R}$ *такая, что* $h(x) = F(f(x))\ \forall x \in X$, *и иначе равна бесконечности.*

2. *В качестве характеристической функции* ω_π *можно взять*

$$\omega_\pi(h; f) = \min_{F:\mathbb{R}\to\mathbb{R}} \Big[\sum_{x \in X} u(h(x), F(f(x))) \Big],$$

где u — функция штрафа на парах яркостей: $u(i,j)=0$, если $i=j$, и равна бесконечности в противном случае.

Доказательство. 1. В соответствии с формулой (1.1), оператор P_f строит проекцию изображения g на множество $V(f)$, т. е. выполняется равенство:

$$\|P_f g - g\|^2 = \min(\|h-g\|^2 \mid h \in V(f)).$$

Обозначим $P_f g \equiv f'$, и перепишем в другом виде:

$$f' = \arg\min(\|h-g\|^2 \mid h \in V(f)).$$

При этом принадлежность $h \in V(f)$ в терминах морфологического анализа означает, что существует функция преобразования яркости $F:\mathbb{R}\to\mathbb{R}$ такая, что $h(x)=F(f(x))\ \forall x \in X$. А это совпадает с условием равенства нулю характеристической функции $\omega_\pi(h,f)$. Тогда h из $V(f)$ можно переписать как $\{h \,|\, h$ — любое, но $\omega_\pi(h;f)=0\}$. Следовательно, проекция f' будет вычисляться по следующей формуле:

$$f' = \arg\min(\|h-g\|^2 \mid h \text{ — любое, но } \omega_\pi(h;f)=0),$$

или, поскольку ω_π принимает только два значения (ноль, бесконечность),

$$f' = \arg\min(\|h-g\|^2 + \omega_\pi(h;f) \mid h \text{ — любое}).$$

2. Поскольку u принимает два значения (ноль, бесконечность), то выражение

$$\min_{F:\mathbb{R}\to\mathbb{R}} \Big[\sum_{x\in X} u(h(x), F(f(x)))\Big]$$

будет равно нулю и достигаться тогда, когда $h(x)=F(f(x))$ для всех точек $x \in X$, т. е. $h \in V(f)$. Таким образом,

$$\min_{F:\mathbb{R}\to\mathbb{R}} \Big[\sum_{x\in X} u(h(x), F(f(x)))\Big] = 0,$$

если существует функция преобразования яркости $F:\mathbb{R}\to\mathbb{R}$ такая, что $h(x)=F(f(x))\ \forall x \in X$, и иначе равна бесконечности, что совпадает с определением характеристической функции $\omega_\pi(h;f)$. ■

В следующей главе (п.3.3.1 и п.3.3.2) аналогичные утверждения будут сформулированы для классов линейных и квадратичных функций.

Пункт 2 теоремы 2.1 дает возможность строить новые алгоритмы путем изменения функции u. Можно задать u конечной, например вида $u(i,j)=\lambda(i-j)^2$. Тогда получим характеристическую функцию структуры

$$\omega(h;f) = \lambda \min_{F:\mathbb{R}\to\mathbb{R}} \Big[\sum_{x\in X} (h(x) - F(f(x)))^2\Big]$$

и оператор преобразования яркостей изображения

$$f' = \arg\min \Big(\|h - g\|^2 + \lambda \min_{F:\mathbb{R}\to\mathbb{R}} \Big[\sum_{x\in X} (h(x) - F(f(x))^2 \Big] \mid h \text{ — любое} \Big).$$

С теоретической точки зрения такой оператор ведет себя более гладко, так как он не имеет скачков по ω, поэтому можно ожидать от него регулярного поведения. Однако на практике вычисление такого оператора — трудоемкая задача, выходящая за рамки данной работы.

2.4 Обнаружение и локализация структурных различий

Введем на пространстве изображений Φ отношение частичного порядка. Для произвольных $f, g \in \Phi$ будем говорить, что структура изображения f не сложнее структуры изображения g, и писать $\pounds_f \preceq \pounds_g$, если для каждого значения i функции f найдется разбиение $L_f(i) = \bigcup\limits_{j=1}^{n} L_g(i_j)$, $n \geq 1$, причем $i_k \neq i_l$ $\forall k, l \in [1, n]$. Если говорить неформально, то на изображении g сохранены все границы изображения f, но к ним, возможно, добавлены новые. Очевидно, что $|\pounds_f| \leq |\pounds_g|$ (здесь $|\cdot|$ — мощность множества).

Утверждение 2.1. *$\pounds_f \preceq \pounds_g$ тогда и только тогда, когда $f = P_g f$.*

Доказательство. При условии $|\pounds_f| = |\pounds_g|$, включающем в себя случай $f = g$, справедливость утверждения очевидна. Ниже полагаем $|\pounds_f| \neq |\pounds_g|$.

1. Пусть $\pounds_f \preceq \pounds_g$. Покажем, что для произвольного $x_0 \in X$ выполняется равенство $f(x_0) = P_g f(x_0)$. Обозначим $f_0 = f(x_0)$ и $g_0 = g(x_0)$. По формуле 1.1

$$P_g f(x_0) = \sum_i \frac{\sum\limits_{x'\in X} f(x') \cdot \chi_i^g(x')}{\sum\limits_{x'\in X} \chi_i^g(x')} \cdot \chi_i^g(x_0) = \frac{1}{|L_g(g_0)|} \sum_{x'\in L_g(g_0)} f(x').$$

По условию $L_g(g_0) \subset L_f(f_0)$, поэтому $\forall x' \in L_g(g_0)$ $f(x') = f_0$ и

$$P_g f(x_0) = \frac{1}{|L_g(g_0)|} \sum_{x'\in L_g(g_0)} f_0 = f_0 = f(x_0),$$

что и требовалось доказать.

2. Покажем, что равенство $f = P_g f$ влечет $\pounds_f \preceq \pounds_g$. Допустим, что это не так, и найдется значение j такое, что $L_g(j) \cap L_f(i_1) \neq \varnothing$ и $L_g(j) \cap L_f(i_2) \neq \varnothing$ для некоторых значений i_1, i_2 функции f. Пусть $x_1 = L_g(j) \cap L_f(i_1)$ и $x_2 = L_g(j) \cap L_f(i_2)$, т.е. $f(x_1) = i_1 \neq i_2 = f(x_2)$ и $g(x_1) = g(x_2) = j$. По определению оператора P_g,

$$P_g f(x_1) = \frac{1}{|L_g(j)|} \sum_{x'\in L_g(j)} f(x') = P_g f(x_2),$$

следовательно, $f(x_1) = P_g f(x_1) = P_g f(x_2) = f(x_2)$, что противоречит выбору точек x_1 и x_2. Полученное противоречие завершает доказательство. ■

Благодаря этому утверждению можно разбить все возможные варианты соотношения структур исходных изображений на следующие три случая.

1. Изображения f и g несравнимы: $\pounds_f \not\preceq \pounds_g$ и $\pounds_g \not\preceq \pounds_f$. Тогда $f \neq P_g f$ и $g \neq P_f g$. В этом случае будем говорить, что структура этих изображений различна, т. е. обнаружено структурное различие изображений в целом (рис. 2.2 (1)).

2. Изображения f и g изоморфны: $\pounds_f \preceq \pounds_g$ и $\pounds_g \preceq \pounds_f$. Тогда $f = P_g f$ и $g = P_f g$. В этом случае будем говорить, что изображения имеют одинаковую структуру, т. е. структурных различий нет (рис. 2.2 (2)).

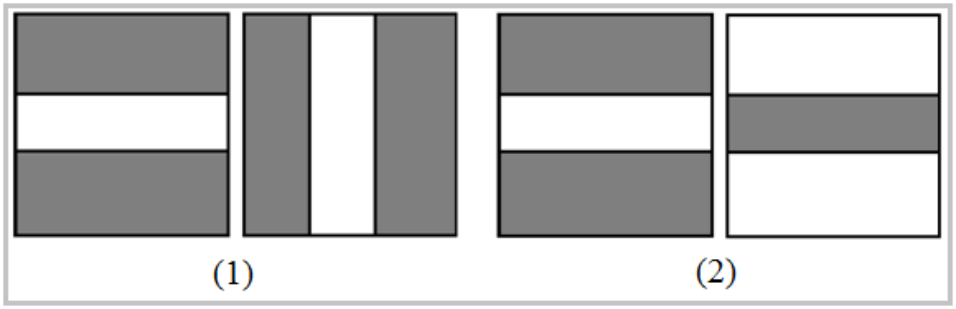

Рис. 2.2. Примеры изображений, чьи структуры несравнимы (1), или изоморфны (2).

3. Для изображений f и g выполнено либо $\pounds_f \preceq \pounds_g$ и $\pounds_g \not\preceq \pounds_f$, либо $\pounds_f \not\preceq \pounds_g$ и $\pounds_g \preceq \pounds_f$. Поскольку эти случаи симметричны, то, не нарушая общности, далее везде будем считать, что выполняются условия $\pounds_f \preceq \pounds_g$ и $\pounds_g \not\preceq \pounds_f$ ($f = P_g f$ и $g \neq P_f g$). В отличии от случаев несравнимых и изоморфных изображений здесь можно говорить не только об обнаружении структурного различия, но и о его локализации, т. е. об определении наличия или отсутствия структурного различия в произвольной зафиксированной точке x_c.

Для исследования этого случая в следующем пункте будет введена математическая модель структурного различия и поставлена задача локализации такого различия.

2.5 Математическая модель структурных различий и постановка задачи

Пусть дана пара изображений $f^0, g^0 \in \Phi$. Первое изображение f^0 (будем называть его фоном) содержит N_f уровней яркости. Значение яркости i-го уровня обозначим f_i^0, площадь уровня — $S_i = |L_{f^0}(f_i^0)|$ ($i \in [1, N_f]$). Изображение g^0 строится путем добавления к i-ому уров-

ню фона объекта, состоящего из N_g^i уровней яркости. Значение яркостей уровней разбиения $j \in [0, N_g^i]$ обозначим $g_{i,j}^0$, $S_{i,j} = |L_{g^0}(g_{i,j}^0)|$, $S_i = \sum\limits_{j=0}^{N_g^i} S_{i,j}$ (индексы j от 1 до N_g^i соответствуют объекту, 0 — фону). Таким образом, $L_{f^0}(f_i^0) = \bigcup\limits_{j=0}^{N_g^i} L_{g^0}(g_{i,j}^0)$, $i \in [1, N_f]$. Чтобы избежать краевых эффектов, добавим условие, что площадь добавленного объекта меньше площади фона. Таким образом, изображение g состоит из $N_f + \sum\limits_{i=1}^{N_f} N_g^i$ уровней яркости, где первое слагаемое соответствует количеству значений функции f^0, второе слагаемое соответствует добавленному объекту. Будем считать, что *в данной точке есть структурное различие*, если эта точка принадлежит объекту. Для моделирования реальной ситуации в задаче поиска структурных различий к каждому изображению добавляем дискретный, стационарный и независимый в каждой точке аддитивный шум с известными параметрами: для первого изображения — n_η, для второго — n_ξ. Здесь n — означает случайное поле, η и ξ — плотности распределения случайных величин, возникающих при фиксировании аргумента n. В соответствии с условиями, налагаемыми на добавляемый шум, случайные поля n_η и n_ξ являются действительными, стационарными (однородными) и удовлетворяют условию эргодичности [42].

Обозначим $f(\cdot) = f^0(\cdot) + n_\eta(\cdot)$ и $g(\cdot) = g^0(\cdot) + n_\xi(\cdot)$. Таким образом, в качестве модели структурного различия используется дискретный (по аргументу и по значению) случайный процесс, который представляет собой сумму неслучайной функции–изображения и независимого дискретного случайного процесса с законом распределения в каждой точке, заданным таблицей:

$$\begin{aligned} f &= \{f(x_i)\}_{i=1}^S = \{f^0(x_i) + n_\eta\}_{i=1}^S, \\ g &= \{g(x_i)\}_{i=1}^S = \{g^0(x_i) + n_\xi\}_{i=1}^S. \end{aligned} \tag{2.2}$$

Случайные поля f и g являются действительными, но, в общем случае, условиям однородности и эргодичности не удовлетворяют, т. к. по причине наличия неслучайной и различной в точках каждого уровня яркости добавки (исходные уровни f^0 и g^0) их мат.ожидания уже не будут константами. Тем не менее, их можно интерпретировать как модели случайных изображений, к которым будем применять алгоритм поиска структурных различий на основе морфологического проектора Пытьева.

Этот алгоритм подробно описан в следующей главе. Дадим здесь его краткое описание. Алгоритм получает на вход пару изображений одинакового размера и состоит из трех шагов. Первый шаг — построение проекции $P_f g$ второго изображения g на форму первого изображения $V(f)$ и проекции $P_g f$ первого изображения f на форму второго изображения $V(g)$. Второй

шаг — построение разностных изображений $R_{fg} = |P_f g - g|$ и $R_{gf} = |P_g f - f|$, после чего строится результирующее разностное изображение $R = \max(R_{fg}, R_{gf})$ (все операции понимаются в поточечном смысле). Третий шаг — пороговая обработка изображения R, параметр порога будем обозначать T. В качестве результатов пороговой обработки можно рассматривать следующие два случая.

1. Если значение функции R в некоторой произвольной зафиксированной точке x_c превосходит заданный порог, то будем говорить, что в этой точке обнаружено структурное различие. В этом случае можно говорить о *локализации* структурного различия в точке x_c.

2. Если количество точек, в которых обнаружено структурное различие, больше некоторого заданного значения, то будем считать, что структура входных изображений различна. В этом случае можно говорить об *обнаружении структурного* различия этих изображений в целом.

Выбор интерпретации обнаружение / локализация зависит от конкретных требований практического применения алгоритма. Такая модельная ситуация достаточно точно отражает возникновение реальных структурных различий, связанных с появлением или исчезновением некоторых объектов (например, домов), и потому представляет определенный практический интерес. На рис. 2.3 представлен пример изображения, полученного путем добавления к однотонному фону объекта, состоящего из трех уровней яркости. Можно видеть, что даже такой простой пример адекватно отражает ситуацию появления на изображении дома с четырехскатной крышей. Тем не менее, реальные изображения намного сложнее, поэтому в четвертой главе диссертации для эмпирического исследования алгоритмов будет предложено развитие данной модели, более близкое к практике, но при этом неподходящее для теоретического исследования.

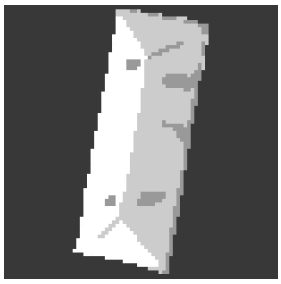

Рис. 2.3. Пример модельного изображения с одноуровневым фоном и трехуровневым объектом, соответствующим появившемуся дому.

Поскольку изображения f и g — случайные поля, то значение разностного изображения $R(x_c)$ в точке $x_c \in X$ также является величиной случайной, и основной задачей будет построение способа ее вычисления. Решение этой задачи позволит определить оптимальное значение порога для алгоритма и оценки ошибок для него. В качестве критерия оптимальности будем

использовать минимум суммы вероятностей двух событий — пропуска структурного различия и ложного обнаружения различия (ошибки первого и второго рода).

Однако значение случайной величины $R(x_c)$ существенно зависит от того, есть структурное различие в точке x_c или его нет, поэтому необходимо рассматривать соответствующие два условных распределения.

У с л о в и е 1: точка x_c принадлежит объекту, т. е. в ней есть структурное различие.

У с л о в и е 2: точка x_c принадлежит фону, т. е. структурного различия в этой точке нет.

Распределение вероятности $r(R(x_c) = i \mid x_c$ принадлежит объекту) обозначим $r^+(i)$. Второе распределение $r(R(x_c) = i \mid x_c$ принадлежит фону) обозначим $r^-(i)$.

З а д а ч а 1. Получить условные распределения вероятностей $r^+(i)$ и $r^-(i)$ величины $R(x_c)$ для модельной задачи.

Для простоты изложения удобно свести решение этой задачи к решению последовательности более простых задач с нарастающей сложностью модели изображения g.

З а д а ч а 1.1. Получить распределения вероятностей $r^+(i)$ и $r^-(i)$ для модельной задачи с параметрами $N_f = 1$, $N_g^1 = 1$.

З а д а ч а 1.2. Получить распределения вероятностей $r^+(i)$ и $r^-(i)$ для модельной задачи с параметрами $N_f = 1$, $N_g^1 = 2$.

З а д а ч а 1.3. Получить распределения вероятностей $r^+(i)$ и $r^-(i)$ для модельной задачи с параметрами $N_f = 1$, $N_g^1 \geq 1$.

Случай $N_f \geq 1$, $N_g^1 \geq 1$ соответствует задаче 1, решение которой позволит определить важный для практических приложений способ вычисления оптимального порога алгоритма поиска структурных различий.

В следующем пункте представлено решение задачи 1 для алгоритма на основе морфологического проектора и применение полученных утверждений для определение оптимальных параметров алгоритма.

2.6 Решение задачи локализации структурных различий

Поскольку алгоритм локализации структурных различий оценивает только одну зафиксированную точку $x_c \in X$, то при использовании морфологического проектора важно только множество $L_f(f(x_c))$, которому эта точка принадлежит. Пусть $k = |L_f(f(x_c))|$. Очевидно, что $k \in [1, S]$. Введем обозначение $L_f^{i,j}(f(x_c)) = \{x \in L_f(f(x_c)) \mid f^0(x) = f_i^0$ и $g^0(x) = g_{i,j}^0\}$. Заметим, что

$$|L_f^{i,j}(f(x_c))| \leq S_{i,j}. \tag{2.3}$$

По построению $L_f^{i,j}(f(x_c)) \subset L_f(f(x_c))$ и $L_f(f(x_c)) = \bigcup\limits_{i,j(i)} L_f^{i,j}(f(x_c))$.

Далее везде под $B(n,m,p)$ будем понимать биномиальное распределение вероятностей: вероятность появления события ровно n раз в серии из m испытаний при условии, что в единичном испытании вероятность его появления равна p [15]. Докажем следующую вспомогательную лемму.

Лемма 2.1. *Пусть дано случайное поле f, заданное по формуле* (2.2).

1. *В задаче* 1.1 *при условии* 1 *вероятность появления реализации поля f такой, что* $|L_f(f(x_c))| = k$, $|L_f^{1,1}(f(x_c))| = a_1 + 1$ *и* $|L_f^{1,0}(f(x_c))| = a_0$, *равна*

$$p_{k,a}^1 = \sum_i \Big(p(f=i) \cdot B(a_1, S_{1,1}-1, p(f=i)) \cdot B(a_0, S_{1,0}, p(f=i)) \Big),$$

где $a = (a_0, a_1)$, $a_0 = k - a_1 - 1$, $a_1 \in [\max\{0, k-1-S_{1,0}\}, \min\{k-1, S_{1,1}-1\}]$.

2. *В задаче* 1.1 *при условии* 2 *вероятность появления реализации поля f такой, что* $|L_f(f(x_c))| = k$, $|L_f^{1,1}(f(x_c))| = a_1$ *и* $|L_f^{1,0}(f(x_c))| = a_0 + 1$, *равна*

$$p_{k,a}^0 = \sum_i \Big(p(f=i) \cdot B(a_1, S_{1,1}, p(f=i)) \cdot B(a_0, S_{1,0}-1, p(f=i)) \Big),$$

где $a = (a_0, a_1)$, $a_0 = k - a_1 - 1$, $a_1 \in [\max(0, k-S_{1,0}), \min(k-1, S_{1,1})]$.

Доказательство. 1. Элементы множества $L_f(f(x_c))$ можно разделить на три группы: $L_f(f(x_c)) = \{x_c\} \cup \{L_f^{1,1}(f(x_c)) \setminus \{x_c\}\} \cup L_f^{1,0}(f(x_c))$. Пусть $|L_f^{1,1}(f(x_c)) \setminus \{x_c\}| = a_1$ и $|L_f^{1,0}(f(x_c))| = a_0$. Таким образом, с учетом (2.3), должна выполняться следующая система:

$$\left.\begin{aligned} &a_0 + a_1 = k-1, \\ &0 \le a_0 \le S_{1,0}, \\ &0 \le a_1 \le S_{1,1} - 1. \end{aligned}\right\}$$

Ограничения можно переписать в следующем виде: $a_1 \in [\max\{0, k-1-S_{1,0}\}, \min\{k-1, S_{1,1}-1\}]$, $a_0 = k - a_1 - 1$.

Занумеруем все точки поля зрения, кроме x_c, числами от 1 до $S-1$: $x_1, x_2, \ldots, x_{S-1}$, причем множество $L_f(f(x_c)) = \{x_1, x_2, \ldots, x_{k-1}, x_c\}$. Зафиксируем элементарное событие i (исход для случайной величины $f(x_j)$). Далее вычислим вероятность появления реализации поля f с таким множеством $L_f(f(x_c))$. Учитывая, что случайные величины $f(x_j)$ независимы, получим

$$p_{k,a}^{(1)} = p(f(x_c) = i) \cdot \prod_{j=1}^{k-1} p(f(x_j) = i) \cdot \prod_{j=k}^{S-1} p(f(x_j) \ne i).$$

Выбирая элементы $L_f(f(x_c))$, мы не могли распоряжаться только зафиксированной точкой x_c. Набор остальных $k-1$ элементов можно выбрать C_{S-1}^{k-1} различными способами. Поэтому можно определить вероятность появления такой реализации поля f, для которой множество $L_f(f(x_c))$ состоит из произвольных $k-1$ элементов и точки x_c, причем $\forall x \in L_f(f(x_c))\ f(x) = i$:

$$p_{k,a}^{(2)} = p(f(x_c) = i) \cdot C_{S-1}^{k-1} \cdot \prod_{j=1}^{k-1} p(f(x_j) = i) \cdot \prod_{j=k}^{S-1} p(f(x_j) \neq i).$$

По условию случайные величины $f(x_j) = f^0(x_j) + n_\eta(x_j)$ распределены одинаково для всех x_j (включая x_c), поэтому нумерация x_j не важна. Имеем

$$p_{k,a}^{(2)} = p(f = i) \cdot C_{S-1}^{k-1} \cdot p^{k-1}(f = i) \cdot p^{S-k}(f \neq i).$$

Чтобы получить вероятность появления реализации f, для которой множество $L_f(f(x_c))$ состоит из произвольных $k-1$ элементов и зафиксированной точки x_c, нужно просуммировать вероятности $p_{k,a}^{(2)}$ по всем возможным исходам i:

$$p_{k,a}^{(3)} = \sum_i \Big(p(f = i) \cdot C_{S-1}^{k-1} \cdot p^{k-1}(f = i) \cdot p^{S-k}(f \neq i) \Big).$$

Теперь уточним способ выбора $k-1$ элемента из $S-1$ точек множества X. Выбирая эти $k-1$ элементов, мы должны разделить их на два подмножества: $L_f^{1,1}(f(x_c)) \setminus \{x_c\}$ и $L_f^{1,0}(f(x_c))$, т. е. мы выбираем a_1 элементов из $S_{1,1}-1$ точек и a_0 элементов из $S_{1,0}$ точек. Таким образом, вместо C_{S-1}^{k-1} способов будет $C_{S_{1,1}-1}^{a_1} \cdot C_{S_{1,0}}^{a_0}$ способ выбора элементов множества $L_f(f(x_c))$:

$$p_{k,a}^{(4)} = \sum_i \Big(p(f = i) \cdot C_{S_{1,1}-1}^{a_1} \cdot C_{S_{1,0}}^{a_0} \cdot p^{k-1}(f = i) \cdot p^{S-k}(f \neq i) \Big).$$

Применяя преобразования

$$p^{S-k}(f \neq i) = p^{(S_{1,0}+S_{1,1})-(1+a_0+a_1)}(f \neq i) = p^{(S_{1,1}-1)-a_1}(f \neq i) \cdot p^{S_{1,0}-a_0}(f \neq i),$$

получим

$$p_{k,a}^{(4)} = \sum_i \Big(p(f = i) \cdot C_{S_{1,1}-1}^{a_1} \cdot p^{a_1}(f = i) \cdot p^{(S_{1,1}-1)-a_1}(f \neq i) \cdot C_{S_{1,0}}^{a_0} \cdot p^{a_0}(f = i) \cdot p^{S_{1,0}-a_0}(f \neq i) \Big).$$

Далее, учитывая, что $p(f \neq i) = 1 - p(f = i)$ и $C_m^x \cdot p^x \cdot (1-p)^{m-x} = B(x, m, p)$ — биномиальное распределение вероятностей, получаем итоговую формулу

$$p_{k,a}^{1} = \sum_i \Big(p(f = i) \cdot B(a_1, S_{1,1}-1, p(f = i)) \cdot B(a_0, S_{1,0}, p(f = i)) \Big).$$

2. Д о к а з а т е л ь с т в о аналогично доказательству п. 1.

Лемма доказана. ■

Теперь перейдем к более сложной задаче 1.2, когда добавленный объект состоит из двух уровней яркости. Введем обозначение $a = (a_0, a_1, a_2)$.

Лемма 2.2. *Пусть дано случайное поле f, заданное по формуле* (2.2).

1. *В задаче* 1.2 *при выполнении условий* 1 *и* $x_c \in L_f^{1,1}(f(x_c))$ *вероятность появления реализации поля f такой, что* $|L_f(f(x_c))| = k$, $|L_f^{1,1}(f(x_c))| = a_1 + 1$, $|L_f^{1,2}(f(x_c))| = a_2$ *и* $|L_f^{1,0}(f(x_c))| = a_0$, *равна*

$$p_{k,a}^1 = \sum_i \Big(p(f=i) \cdot B(a_1, S_{1,1} - 1, p(f=i)) \cdot B(a_2, S_{1,2}, p(f=i)) \cdot B(a_0, S_{1,0}, p(f=i)) \Big),$$

где

$$\left.\begin{aligned} &a_0 + a_1 + a_2 = k - 1, \\ &0 \le a_0 \le S_{1,0}, \\ &0 \le a_1 \le S_{1,1} - 1, \\ &0 \le a_2 \le S_{1,2}. \end{aligned}\right\}$$

Если же $x_c \in L_f^{1,2}(f(x_c))$, *то* $|L_f^{1,1}(f(x_c))| = a_1$, $|L_f^{1,2}(f(x_c))| = a_2 + 1$, $|L_f^{1,0}(f(x_c))| = a_0$, *и эта вероятность равна*

$$p_{k,a}^2 = \sum_i \Big(p(f=i) \cdot B(a_1, S_{1,1}, p(f=i)) \cdot B(a_2, S_{1,2} - 1, p(f=i)) \cdot B(a_0, S_{1,0}, p(f=i)) \Big),$$

где

$$\left.\begin{aligned} &a_0 + a_1 + a_2 = k - 1, \\ &0 \le a_0 \le S_{1,0}, \\ &0 \le a_1 \le S_{1,1}, \\ &0 \le a_2 \le S_{1,2} - 1. \end{aligned}\right\}$$

2. *В задаче* 1.2 *при условии* 2 *вероятность появления реализации поля f такой, что* $|L_f(f(x_c))| = k$, $|L_f^{1,1}(f(x_c))| = a_1$, $|L_f^{1,2}(f(x_c))| = a_2$ *и* $|L_f^{1,0}(f(x_c))| = a_0 + 1$, *равна*

$$p_{k,a}^0 = \sum_i \Big(p(f=i) \cdot B(a_1, S_{1,1}, p(f=i)) \cdot B(a_2, S_{1,2}, p(f=i)) \cdot B(a_0, S_{1,0} - 1, p(f=i)) \Big),$$

где

$$\left.\begin{aligned} &a_0 + a_1 + a_2 = k - 1, \\ &0 \le a_0 \le S_{1,0} - 1, \\ &0 \le a_1 \le S_{1,1}, \\ &0 \le a_2 \le S_{1,2}. \end{aligned}\right\}$$

Д о к а з а т е л ь с т в о. Как и в лемме 2.1, разобьем все возможные реализации поля f на классы в зависимости от принадлежности зафиксированной точки x_c определенному подмножеству $L_f(f(x_c))$: $L_f^{1,0}(f(x_c))$ (вероятность $p_{k,a}^0$), $L_f^{1,1}(f(x_c))$ (вероятность $p_{k,a}^1$) или $L_f^{1,2}(f(x_c))$ (вероятность $p_{k,a}^2$).

1. Найдем вероятность $p_{k,a}^1$. В отличии от леммы 2.1 набор элементов, составляющих множество $L_f(f(x_c))$, делится уже на четыре группы. Первая группа — это зафиксированная точка x_c. Вторая и третья группы — элементы множеств $L_f^{1,1}(f(x_c))\setminus\{x_c\}$ и $L_f^{1,2}(f(x_c))$. Мощность этих множеств $|L_f^{1,1}(f(x_c))\setminus\{x_c\}| = a_1$ и $|L_f^{1,2}(f(x_c))| = a_2$. Четвертая группа — оставшиеся точки $L_f^{1,0}(f(x_c))$ ($|L_f^{1,0}(f(x_c))| = a_0$). Таким образом, должна выполняться следующая система:

$$\left.\begin{aligned} &a_0 + a_1 + a_2 = k-1,\\ &0 \le a_0 \le S_{1,0},\\ &0 \le a_1 \le S_{1,1}-1,\\ &0 \le a_2 \le S_{1,2}, \end{aligned}\right\}$$

где $k = 1 + a_0 + a_1 + a_2$. В этой системе необходимость равенства очевидна, а неравенства следуют из того, что $a_0 = |L_f^{1,0}(f(x_c))| \le S_{1,0}$, $a_1 = |L_f^{1,1}(f(x_c))\setminus\{x_c\}| = |L_f^{1,1}(f(x_c))| - 1 \le S_{1,1}-1$ и $a_2 = |L_f^{1,2}(f(x_c))| \le S_{1,2}$.

В лемме 2.1 была получена вероятность появления такой реализации f, для которой множество $L_f(f(x_c))$ состоит из произвольно выбранных $k-1$ элементов из $S-1$ точек поля зрения X и зафиксированного элемента x_c. Имеем

$$p_{k,a}^{(1)} = \sum_i \big(p(f=i)\cdot C_{S-1}^{k-1}\cdot p^{k-1}(f=i)\cdot p^{S-k}(f\ne i)\big).$$

В процессе выбора $k-1$ элементов происходит разделение их на три подмножества: $L_f^{1,1}(f(x_c))\setminus\{x_c\}$, $L_f^{1,2}(f(x_c))$ и $L_f^{1,0}(f(x_c))$, т. е. выбираются a_1 элементов из $S_{1,1}-1$ точек, a_2 — из $S_{1,2}$ точек и a_0 — из $S_{1,0}$ точек. Таким образом, вместо C_{S-1}^{k-1} способов будет $C_{S_{1,1}-1}^{a_1}\cdot C_{S_{1,2}}^{a_2}\cdot C_{S_{1,0}}^{a_0}$ способ выбора элементов множества $L_f(f(x_c))$. Получим

$$p_{k,a}^{(2)} = \sum_i \Big(p(f=i)\cdot C_{S_{1,1}-1}^{a_1}\cdot C_{S_{1,2}}^{a_2}\cdot C_{S_{1,0}}^{a_0}\cdot p^{k-1}(f=i)\cdot p^{S-k}(f\ne i)\Big).$$

После проведения следующих преобразований

$$p^{S-k}(f\ne i) = p^{(S_{1,0}+S_{1,1}+S_{1,2})-(1+a_0+a_1+a_2)}(f\ne i)$$

$$= p^{(S_{1,1}-1)-a_1}(f\ne i)\cdot p^{S_{1,2}-a_2}(f\ne i)\cdot p^{S_{1,0}-a_0}(f\ne i)$$

имеем

$$p_{k,a}^{(2)} = \sum_i \Big(p(f=i)\cdot C_{S_{1,1}-1}^{a_1}\cdot p^{a_1}(f=i)\cdot p^{(S_{1,1}-1)-a_1}(f\ne i)$$

$$\times C_{S_{1,2}}^{a_2} \cdot p^{a_2}(f=i) \cdot p^{S_{1,2}-a_2}(f \neq i)) \cdot C_{S_{1,0}}^{a_0} \cdot p^{a_0}(f=i) \cdot p^{S_{1,0}-a_0}(f \neq i)\Big).$$

Далее, учитывая, что $p(f \neq i) = 1 - p(f=i)$ и $C_m^x \cdot p^x \cdot (1-p)^{m-x} = B(x, m, p)$, получаем итоговую формулу

$$p_{k,a}^1 = \sum_i \Big(p(f=i) \cdot B(a_1, S_{1,1} - 1, p(f=i)) \cdot B(a_2, S_{1,2}, p(f=i)) \cdot B(a_0, S_{1,0}, p(f=i)) \Big).$$

Вероятность $p_{k,a}^2$ вычисляется аналогично вероятности $p_{k,a}^1$.

2. Доказательство аналогично доказательствам леммы 2.1 и п. 1.

Лемма доказана. ■

Подобным образом можно доказать и более общий факт, когда объект на втором изображении состоит из N_g^1 уровней яркости (задача 1.3). Теперь можно приступить к рассмотрению самого общего случая, когда и фон, и объект состоят из нескольких уровней яркости. Введем обозначения: конфигурация $a = (a_{1,0}, a_{1,1}, \ldots, a_{1,N_g^1}, \ldots, a_{N_f,0}, a_{N_f,1}, \ldots, a_{N_f,N_g^{N_f}})$, где $a_{q,j}$ соответствует мощности пересечения множества $L_f(f(x_c))$ и уровня яркости объекта с номером j, полученного из разбиения уровня фона с номером q (индекс 0 соответствует фону на втором изображении). Из-за разницы яркостей незашумленного изображения f^0 в точках разных уровней после добавления шума распределения вероятностей яркостей в этих уровнях также будут различными, поэтому далее будем обозначать $p_q(f=i) = p(f(x)=i \mid f^0(x) = f_q^0)$ — вероятность появления яркости i в точке, которая на изображении f^0 принадлежит уровню яркости с номером q.

Теорема 2.2. *Пусть дано случайное поле f, заданное по формуле* (2.2).

1. *В задаче* 1 *при выполнении условий* 1 *и* $x_c \in L_f^{n,m}(f(x_c))$ $(n \in [1, N_f]$ *и* $m \in [1, N_g^n])$ *вероятность появления реализации поля f такой, что* $|L_f(f(x_c))| = k$, $|L_f^{n,m}(f(x_c))| = a_{n,m} + 1$ *и* $|L_f^{q,j}(f(x_c))| = a_{q,j}$ *для* (q,j): $q \in [1, N_f]$, $j \in [0, N_g^q]$, $(q,j) \neq (n,m)$, *равна*

$$p_{k,a}^{n,m} = \sum_i \Bigg(p_n(f=i) \cdot \prod_{q=1}^{n-1} \prod_{j=0}^{N_g^q} B(a_{q,j}, S_{q,j}, p_q(f=i))$$

$$\times \prod_{j=0}^{m-1} B(a_{n,j}, S_{n,j}, p_n(f=i)) \cdot B(a_{n,m}, S_{n,m} - 1, p_n(f=i)) \cdot \prod_{j=m+1}^{N_g^n} B(a_{n,j}, S_{n,j}, p_n(f=i))$$

$$\times \prod_{q=n+1}^{N_f} \prod_{j=0}^{N_g^q} B(a_{q,j}, S_{q,j}, p_q(f=i)) \Bigg),$$

где

$$\left.\begin{aligned}&\sum_{q=1}^{N_f}\sum_{j=1}^{N_g^q} a_{q,j} = k-1,\\&0 \le a_{q,j} \le S_{q,j} \quad \forall (q,j)\colon q\in[1,N_f],\ j\in[0,N_g^q], \quad (q,j)\ne(n,m),\\&0 \le a_{n,m} \le S_{n,m}-1.\end{aligned}\right\}$$

2. *В задаче* 1 *при выполнении условий* 2 *и* $x_c \in L_f^{n,0}(f(x_c))$ $(n \in [1,N_f])$ *вероятность появления реализации поля* f *такой, что* $|L_f(f(x_c))| = k$, $|L_f^{n,0}(f(x_c))| = a_{n,0}+1$ *и* $|L_f^{q,j}(f(x_c))| = a_{q,j}$ *для* $(q,j)\colon q\in[1,N_f]$, $j\in[0,N_g^q]$, $(q,j)\ne(n,0)$, *равна*

$$p_{k,a}^{n,0} = \sum_i \Bigg(p_n(f=i)\cdot \prod_{q=1}^{n-1}\prod_{j=0}^{N_g^q} B(a_{q,j},S_{q,j},p_q(f=i))$$

$$\times B(a_{n,0},S_{n,0}-1,p_n(f=i))\cdot \prod_{j=1}^{N_g^n} B(a_{n,j},S_{n,j},p_n(f=i))\cdot \prod_{q=n+1}^{N_f}\prod_{j=0}^{N_g^q} B(a_{q,j},S_{q,j},p_q(f=i))\Bigg),$$

где

$$\left.\begin{aligned}&\sum_{q=1}^{N_f}\sum_{j=1}^{N_g^q} a_{q,j} = k-1,\\&0 \le a_{n,0} \le S_{n,0}-1,\\&0 \le a_{q,j} \le S_{q,j}, \quad \forall (q,j)\colon q\in[1,N_f],\ j\in[0,N_g^q], \quad (q,j)\ne(n,0).\end{aligned}\right\}$$

Доказательство. 1. Элементы множества $L_f(f(x_c))$ можно разбить на $N_f + \sum_{q=1}^{N_f} N_g^q + 1$ подмножеств следующим образом: $L_f(f(x_c)) = \{x_c\} \cup \{L_f^{n,m}(f(x_c)) \setminus \{x_c\}\} \cup \{\bigcup_{(q,j)} L_f^{q,j}(f(x_c))\}$, где $(q,j)\colon q\in[1,N_f]$, $j\in[0,N_g^q]$, $(q,j)\ne(n,m)$. Пусть $|L_f^{n,m}(f(x_c))| = a_{n,m}+1$ и $|L_f^{q,j}(f(x_c))| = a_{q,j}$. Таким образом, с учетом (2.3) должна выполняться следующая система:

$$\left.\begin{aligned}&\sum_{q=1}^{N_f}\sum_{j=1}^{N_g^q} a_{q,j} = k-1,\\&0 \le a_{q,j} \le S_{q,j} \quad \forall (q,j)\colon q\in[1,N_f],\ j\in[0,N_g^q], \quad (q,j)\ne(n,m),\\&0 \le a_{n,m} \le S_{n,m}-1.\end{aligned}\right\}$$

Занумеруем все точки поля зрения, кроме x_c, числами от 1 до $S-1$: $x_1, x_2, \ldots, x_{S-1}$, причем множество $L_f(f(x_c)) = \{x_1, x_2, \ldots, x_{k-1}, x_c\}$. Зафиксируем элементарное событие i (исход для случайной величины $f(x_l)$). Далее вычислим вероятность появления реализации поля f с таким множеством $L_f(f(x_c))$. Учитывая, что случайные величины $f(x_l)$ независимы, получим

$$p_{k,a}^{(1)} = p(f(x_c)=i)\cdot \prod_{l=1}^{k-1} p(f(x_l)=i)\cdot \prod_{l=k}^{S-1} p(f(x_l)\ne i).$$

Выбирая элементы $L_f(f(x_c))$, мы не могли распоряжаться только зафиксированной точкой x_c. Набор остальных $k-1$ элементов можно выбрать C_{S-1}^{k-1} различными способами. Поэтому можно определить вероятность появления такой реализации поля f, для которой множество $L_f(f(x_c))$ состоит из произвольных $k-1$ элементов и точки x_c:

$$p_{k,a}^{(2)} = p(f(x_c) = i) \cdot C_{S-1}^{k-1} \cdot \prod_{l=1}^{k-1} p(f(x_l) = i) \cdot \prod_{l=k}^{S-1} p(f(x_l) \neq i).$$

По условию случайные величины $f(x_l) = f^0(x_l) + n_\eta(x_l)$ распределены одинаково для всех $x_l \in L_f^{q,j}(f(x_c))$, $q \in [1, N_f]$, $j \in [0, N_g^q]$ (включая x_c), поэтому нумерация x_l не важна. Имеем

$$p_{k,a}^{(2)} = p_n(f = i) \cdot C_{S-1}^{k-1} \cdot \prod_{q=1}^{N_f} \prod_{j=0}^{N_g^q} p_q^{a_{q,j}}(f = i) \cdot \prod_{q=1}^{N_f} \prod_{j=0}^{N_g^q} p_q^{S_{q,j}-a_{q,j}}(f \neq i).$$

Чтобы получить вероятность появления реализации f, для которой множество $L_f(f(x_c))$ состоит из произвольных $k-1$ элементов и зафиксированной точки x_c, нужно просуммировать вероятности $p_{k,a}^{(2)}$ по всем возможным исходам i:

$$p_{k,a}^{(3)} = \sum_i \left(p_n(f = i) \cdot C_{S-1}^{k-1} \cdot \prod_{q=1}^{N_f} \prod_{j=0}^{N_g^q} p_q^{a_{q,j}}(f = i) \cdot \prod_{q=1}^{N_f} \prod_{j=0}^{N_g^q} p_q^{S_{q,j}-a_{q,j}}(f \neq i) \right).$$

Теперь уточним способ выбора $k-1$ элемента из $S-1$ точек множества X. При выборе $k-1$ элементов происходит разделение их на $N_f + \sum_{q=1}^{N_f} N_g^q$ подмножеств, т. е. мы выбираем $a_{n,m}$ элементов из $S_{n,m}-1$ и $a_{q,j}$ элементов из $S_{q,j}$ точек. Поэтому вместо C_{S-1}^{k-1} способов будет $\prod\limits_{q=1}^{n-1} \prod\limits_{j=0}^{N_g^q} C_{S_{q,j}}^{a_{q,j}} \cdot \prod\limits_{j=0}^{m-1} C_{S_{n,j}}^{a_{n,j}} \cdot C_{S_{n,m}-1}^{a_{n,m}} \cdot \prod\limits_{j=m+1}^{N_g^n} C_{S_{n,j}}^{a_{n,j}} \cdot \prod\limits_{q=n+1}^{N_f} \prod\limits_{j=0}^{N_g^q} C_{S_{q,j}}^{a_{q,j}}$ способов выбора элементов множества $L_f(f(x_c))$:

$$p_{k,a}^{(4)} = \sum_i \Bigg(p_n(f = i) \cdot \prod_{q=1}^{n-1} \prod_{j=0}^{N_g^q} C_{S_{q,j}}^{a_{q,j}} \cdot \prod_{j=0}^{m-1} C_{S_{n,j}}^{a_{n,j}} \cdot C_{S_{n,m}-1}^{a_{n,m}} \cdot \prod_{j=m+1}^{N_g^n} C_{S_{n,j}}^{a_{n,j}} \cdot \prod_{q=n+1}^{N_f} \prod_{j=0}^{N_g^q} C_{S_{q,j}}^{a_{q,j}}$$

$$\times \prod_{q=1}^{N_f} \prod_{j=0}^{N_g^q} p_q^{a_{q,j}}(f = i) \cdot \prod_{q=1}^{N_f} \prod_{j=0}^{N_g^q} p_q^{S_{q,j}-a_{q,j}}(f \neq i) \Bigg).$$

Группируя множители с одинаковыми индексами, получим

$$p_{k,a}^{n,m} = \sum_i \Bigg(p_n(f = i) \cdot \prod_{q=1}^{n-1} \prod_{j=0}^{N_g^q} \left(C_{S_{q,j}}^{a_{q,j}} \cdot p_q^{a_{q,j}}(f = i) \cdot p_q^{S_{q,j}-a_{q,j}}(f \neq i) \right)$$

$$\times \prod_{j=0}^{m-1} \left(C_{S_{n,j}}^{a_{n,j}} \cdot p_n^{a_{n,j}}(f = i) \cdot p_n^{S_{n,j}-a_{n,j}}(f \neq i) \right) \cdot \left(C_{S_{n,m}-1}^{a_{n,m}} \cdot p_n^{a_{n,m}}(f = i) \cdot p_n^{S_{n,m}-1-a_{n,m}}(f \neq i) \right)$$

$$\times \prod_{j=m+1}^{N_g^n} \left(C_{S_{n,j}}^{a_{n,j}} \cdot p_n^{a_{n,j}}(f=i) \cdot p_n^{S_{n,j}-a_{n,j}}(f \neq i)\right) \cdot \prod_{q=n+1}^{N_f} \prod_{j=0}^{N_g^q} \left(C_{S_{q,j}}^{a_{q,j}} \cdot p_q^{a_{q,j}}(f=i) \cdot p_q^{S_{q,j}-a_{q,j}}(f \neq i)\right)\Bigg).$$

Далее, учитывая, что $p(f \neq i) = 1 - p(f = i)$ и $C_m^x\, p^x\, (1-p)^{m-x} = B(x, m, p)$ — биномиальное распределение вероятностей, получаем итоговую формулу

$$p_{k,a}^{n,m} = \sum_i \Bigg(p_n(f=i) \cdot \prod_{q=1}^{n-1} \prod_{j=0}^{N_g^q} B\big(a_{q,j}, S_{q,j}, p_q(f=i)\big)$$

$$\times \prod_{j=0}^{m-1} B\big(a_{n,j}, S_{n,j}, p_n(f=i)\big) \cdot B\big(a_{n,m}, S_{n,m}-1, p_n(f=i)\big) \cdot \prod_{j=m+1}^{N_g^n} B\big(a_{n,j}, S_{n,j}, p_n(f=i)\big)$$

$$\times \prod_{q=n+1}^{N_f} \prod_{j=0}^{N_g^q} B\big(a_{q,j}, S_{q,j}, p_q(f=i)\big)\Bigg).$$

2. Д о к а з а т е л ь с т в о аналогично доказательству п. 1.

Теорема доказана. ■

Таким образом, пространство всех изображений разбивается на классы в зависимости от параметров (k, a) путем перебора всех возможных вариантов реализации поля (изображения) f. Теорема 2.2 позволяет определить вероятности для классов из этого разбиения. Докажем основную теорему.

Теорема 2.3. 1. *В задаче* 1 *при условии* 1 *распределение вероятностей случайной величины* $R(x_c)$ *имеет вид*

$$r^+(i) = \frac{\sum\limits_{n=1}^{N_f} \sum\limits_{m=1}^{N_g^n} (S_{n,m} \cdot r_{n,m}(i))}{\sum\limits_{u=1}^{N_f} \sum\limits_{v=1}^{N_g^u} S_{u,v}},$$

где

$$r_{n,m}(i) = \sum_{k,a} (p_{k,a}^{n,m} \cdot p(|G_{k,a}^{n,m} + E_k| = i)),$$

величина $p_{k,a}^{n,m}$ *определяется теоремой* 2.2,

$$G_{k,a}^{n,m} = \frac{1}{k} \left(\sum_{q=1}^{N_f} \sum_{j=0}^{N_g^q} (a_{q,j} \cdot g_{q,j}^0) + (1-k)\, g_{n,m}^0 \right), \quad E_k = \frac{1}{k} \sum_{j=1}^{k-1} \varepsilon_j + \frac{1-k}{k}\, \varepsilon_c,$$

$\varepsilon_j = n_\xi(x_j)$ — *случайная величина, элемент случайного поля в точке* $x_j \in L_f(f(x_c))$,
$\varepsilon_c = n_\xi(x_c)$ — *случайная величина, соответствующая точке* x_c.

2. *В задаче* 1 *при условии* 2 *распределение вероятностей случайной величины* $R(x_c)$ *имеет вид*

$$r^-(i) = \frac{\sum\limits_{n=1}^{N_f}(S_{n,0}\cdot r_{n,0}(i))}{\sum\limits_{u=1}^{N_f} S_{u,0}},$$

где

$$r_{n,0}(i) = \sum_{k,a}(p_{k,a}^{n,0}\cdot p(|G_{k,a}^{n,0}+E_k|=i)),$$

величина $p_{k,a}^{n,0}$ *определяется теоремой* 2.2,

$$G_{k,a}^{n,0} = \frac{1}{k}\left(\sum_{q=1}^{N_f}\sum_{j=0}^{N_g^q}(a_{q,j}\cdot g_{q,j}^0) + (1-k)\,g_{n,0}^0\right),$$

определения E_k, ε_j *и* ε_c *такие же, как в п.* 1.

Доказательство. 1. Для реализации поля f с параметрами (k,a) и случайного поля g вычислим значение величины $R(x_c)$ при условии, что $x_c \in L_f^{n,m}(f(x_c))$ (обозначим эту величину $R_{n,m}(x_c)$):

$$R_{n,m}(x_c) = |P_f g(x_c) - g(x_c)| = \left|\frac{1}{k}\sum_{j=1}^{k} g(x_j^{k,a}) - g(x_c)\right|$$

$$= \left|\frac{1}{k}\left(\sum_{q=1}^{N_f}\sum_{j=0}^{N_g^q}(a_{q,j}\cdot g_{q,j}^0) + g_{n,m}^0 + \sum_{j=1}^{k}\varepsilon_j\right) - (g_{n,m}^0+\varepsilon_c)\right|$$

$$= \left|\frac{1}{k}\left(\sum_{q=1}^{N_f}\sum_{j=0}^{N_g^q}(a_{q,j}\cdot g_{q,j}^0) + (1-k)\,g_{n,m}^0\right) + \frac{1}{k}\sum_{j=1}^{k-1}\varepsilon_j + \frac{1-k}{k}\,\varepsilon_c\right|.$$

Здесь

$$G_{k,a}^{n,m} = \frac{1}{k}\left(\sum_{q=1}^{N_f}\sum_{j=0}^{N_g^q}(a_{q,j}\cdot g_{q,j}^0) + (1-k)\,g_{n,m}^0\right) \text{ — константа при фиксированных}(k,a),$$

$$E_k = \frac{1}{k}\sum_{j=1}^{k-1}\varepsilon_j + \frac{1-k}{k}\,\varepsilon_c \text{ — случайная величина,}$$

и, значит, $R_{n,m}(x_c) = |G_{k,a}^{n,m} + E_k|$. Таким образом, каждой вероятности $p_{k,a}^{n,m}$ из теоремы 2.2 соответствует случайная величина $R_{n,m}(x_c)$.

Пусть $R_{n,m}^{\min} = \min\limits_{k,a} R_{n,m}(x_c)$, $R_{n,m}^{\max} = \max\limits_{k,a} R_{n,m}(x_c)$. Составим таблицу 2.1: каждая строка соответствует классу реализаций поля f с параметрами (k,a), каждый столбец — значению

Таблица 2.1

Распределения яркости пикселя разностного изображения в зависимости от параметров (k, a)

(k,a)	$R_{n,m}^{\min}$	$R_{n,m}^{\min}+1$	$\dots$	$R_{n,m}^{\max}-1$	$R_{n,m}^{\max}$
$(1,a)$	$p_{1a}^{n,m}$ $\times p(R(x_c)=R_{n,m}^{\min})$	$p_{1a}^{n,m}$ $\times p(R(x_c)=R_{n,m}^{\min}+1)$		$p_{1a}^{n,m}$ $\times p(R(x_c)=R_{n,m}^{\max}-1)$	$p_{1a}^{n,m}$ $\times p(R(x_c)=R_{n,m}^{\max})$
$(2,a)$	$p_{2a}^{n,m}$ $\times p(R(x_c)=R_{n,m}^{\min})$	$p_{2a}^{n,m}$ $\times p(R(x_c)=R_{n,m}^{\min}+1)$		$p_{2a}^{n,m}$ $\times p(R(x_c)=R_{n,m}^{\max}-1)$	$p_{2a}^{n,m}$ $\times p(R(x_c)=R_{n,m}^{\max})$
$(3,a)$	$p_{3a}^{n,m}$ $\times p(R(x_c)=R_{n,m}^{\min})$	$p_{3a}^{n,m}$ $\times p(R(x_c)=R_{n,m}^{\min}+1)$		$p_{3a}^{n,m}$ $\times p(R(x_c)=R_{n,m}^{\max}-1)$	$p_{3a}^{n,m}$ $\times p(R(x_c)=R_{n,m}^{\max})$
$\dots$					
(S,a)	$p_{Sa}^{n,m}$ $\times p(R(x_c)=R_{n,m}^{\min})$	$p_{Sa}^{n,m}$ $\times p(R(x_c)=R_{n,m}^{\min}+1)$		$p_{Sa}^{n,m}$ $\times p(R(x_c)=R_{n,m}^{\max}-1)$	$p_{Sa}^{n,m}$ $\times p(R(x_c)=R_{n,m}^{\max})$

случайной величины $R_{n,m}(x_c)$. Каждое поле таблицы — совместная реализация этих двух случайных событий, а записанное в это поле значение — вероятность, равная произведению вероятностей исходных событий. Фактически, строки таблицы определяются распределением вероятностей случайных величин, составляющих поле f, столбцы — распределением вероятностей случайных величин, составляющих поле g.

Суммируя таблицу 2.1 по столбцам и переписывая $p_{k,a}^{n,m} \cdot p(R_{n,m}(x_c)=i) = p_{k,a}^{n,m} \cdot p(|G_{k,a}^{n,m} + E_k| = i)$, получаем распределение вероятности случайной величины $R(x_c)$ при условии $x_c \in L_f^{n,m}(f(x_c))$:

$$p\big(R_{n,m}(x_c)=i \mid x_c \in L_f^{n,m}(f(x_c))\big) \triangleq r_{n,m}(i) = \sum_{k,a} \big(p_{k,a}^{n,m} \cdot p(|G_{k,a}^{n,m} + E_k| = i)\big).$$

Далее, применяя формулу полной вероятности

$$r^+(i) = \sum_{n=1}^{N_f}\sum_{m=1}^{N_g^n} \Big(p\big(R_{n,m}(x_c)=i \mid x_c \in L_f^{n,m}(f(x_c))\big) \cdot p\big(x_c \in L_f^{n,m}(f(x_c))\big)\Big)$$

$$= \sum_{n=1}^{N_f}\sum_{m=1}^{N_g^n} \left(r_{n,m}(i) \cdot \frac{S_{n,m}}{\sum\limits_{u=1}^{N_f}\sum\limits_{v=1}^{N_g^u} S_{u,v}} \right) = \frac{\sum\limits_{n=1}^{N_f}\sum\limits_{m=1}^{N_g^n} (S_{n,m} \cdot r_{n,m}(i))}{\sum\limits_{u=1}^{N_f}\sum\limits_{v=1}^{N_g^u} S_{u,v}},$$

получаем искомое выражение.

2. Доказательство аналогично доказательству п. 1. Теорема доказана. ■

Опираясь на теорему 2.3, можно определить способ вычисления оптимального порога алгоритма. Как было указано в пункте 1.4, в качестве критерия оптимальности будет использоваться *percentage correct classification*, т. е. минимум суммы ошибок первого и второго рода. Имеет место следующая теорема.

Теорема 2.4. *В условиях задачи* 1 *порог алгоритма поиска структурных различий на основе морфологического проектора, вычисленный по следующей формуле*

$$T_{opt} = \arg\min_T \left(\sum_{i<T} r^+(i) + \sum_{i>T} r^-(i) \right),$$

является оптимальным в смысле критерия percentage correct classification.

Доказательство. Используя полученные в теореме 2.3 формулы распределений, для фиксированного значения порога T можно выписать формулы для вычисления вероятности пропуска $P_{FN} = \sum\limits_{i<T} r^+(i)$ и вероятности ложной тревоги $P_{FP} = \sum\limits_{i>T} r^-(i)$. Поскольку критерием оптимальности был выбран минимум суммы вероятностей ошибок первого и второго рода, то формула для оптимального порога будет иметь вид

$$T_{opt} = \arg\min_T (P_{FN} + P_{FP}) = \arg\min_T \left(\sum_{i<T} r^+(i) + \sum_{i>T} r^-(i) \right).$$

Теорема доказана. ■

Далее приводятся результаты численного эксперимента по применению теоремы 2.4:

1. На рис. 2.4 приведены графики распределений вероятностей $r^+(i)$ и $r^-(i)$, полученных при условиях задачи 1.1. Параметры изображений $S = 441$, $S_1 = 120$, $g_1^0 = 100$, $g_0^0 = 0$. Оба изображения содержат дискретный шум, распределение которого аппроксимируется нормальным распределением с параметрами $N(0, 10)$.

Выброс в нуле связан с тем, что в случае, когда на первом изображении f уровень яркости, содержащий пиксель x_c, состоит только из этого пикселя (т. е. $k = 1$), то и на втором изображении g будет взята яркость только одного пикселя, и $P_f g(x_c) \equiv g(x_c)$. Яркость этой точки на разностном изображении будет ноль, независимо от добавленного на изображение g шума. Однако с увеличением размера изображений вероятность появления изображения f с параметром $k = 1$ уменьшается, следовательно, уменьшается и величина выброса. Для сравнения приведем графики распределений вероятностей $r^+(i)$ и $r^-(i)$ в случае запрета появления изображений f с $k = 1$ (рис. 2.5). Фактически это означает, что в лемме 2.1 и теореме 2.3 $k \in [2, S]$. На представленных графиках можно видеть, что выброс исчез.

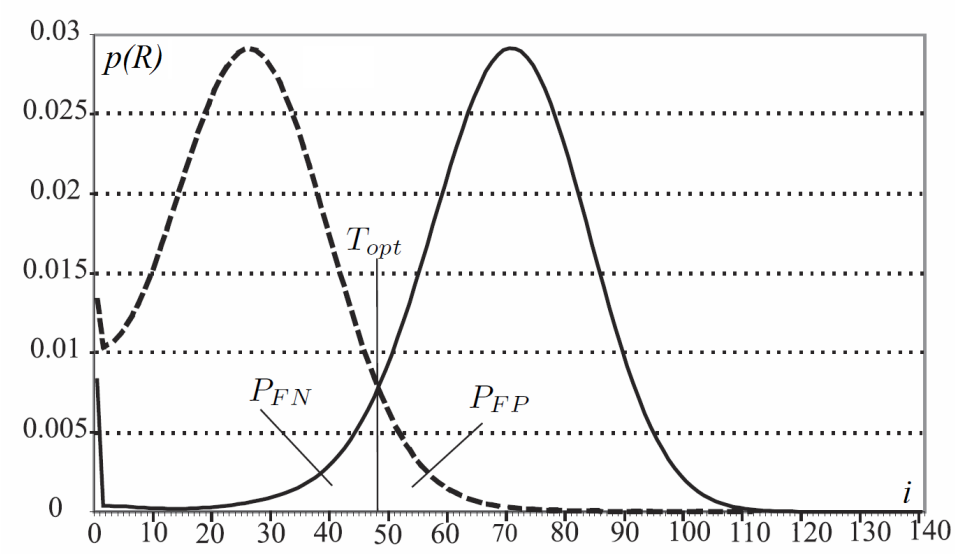

Рис. 2.4. Графики распределений $r^+(i)$ (сплошная линия) и $r^-(i)$ (пунктирная линия) для изображений с параметрами $S = 441$, $S_1 = 120$, $g_1^0 = 100$, $g_0^0 = 0$. Оптимальный порог $T_{opt} = 48$. Оценки ошибок: $P_{FN} = 0.076$, $P_{FP} = 0.049$.

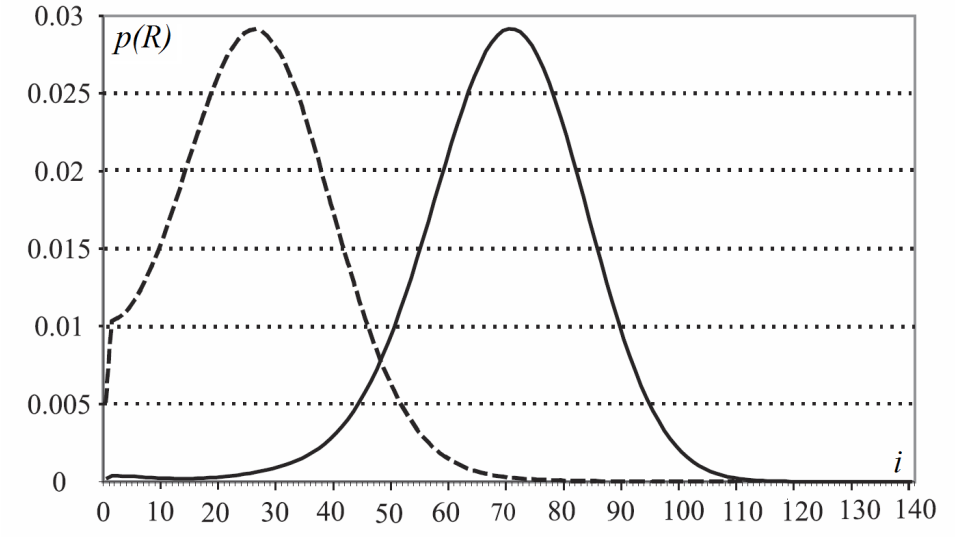

Рис. 2.5. Графики распределений $r^+(i)$ (сплошная линия) и $r^-(i)$ (пунктирная линия) для изображений с параметрами $S = 441$, $S_1 = 120$, $g_1^0 = 100$, $g_0^0 = 0$ и запретом на появление изображения f с параметром $k = 1$.

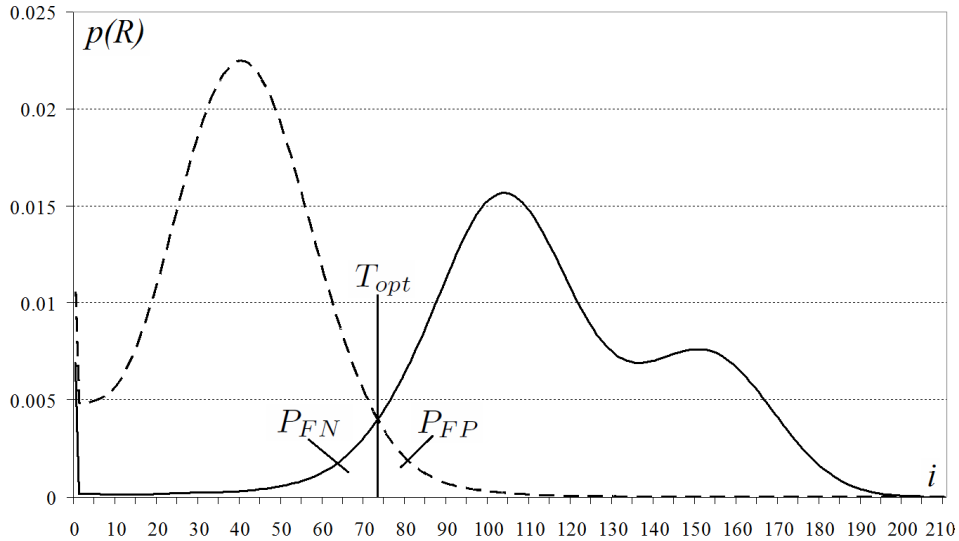

Рис. 2.6. Графики распределений $r^+(i)$ (сплошная линия) и $r^-(i)$ (пунктирная линия) для изображений с параметрами $S = 441$, $S_1 = 40$, $S_2 = 60$, $S_3 = 20$, $g_1^0 = 255$, $g_2^0 = 204$, $g_3^0 = 153$, $g_0^0 = 56$ (добавление трехуровневого объекта). Оптимальный порог $T_{opt} = 73$. Оценки ошибок: $P_{FN} = 0.056$, $P_{FP} = 0.038$.

2. Для сравнения полученных теоретических распределений с эмпирическими генерируем исходные изображения с заданными параметрами, применяем к ним алгоритм поиска структурных различий на основе морфологического проектора и строим распределение для значений яркости центральной точки разностного изображения. Для изображений с параметрами: $S = 2601$, $S_1 = 441$, $g_1^0 = 100$, $g_0^0 = 0$, $n_\eta = N(0, 10)$, $n_\xi = N(0, 10)$ и миллиона сгенерированных изображений средняя ошибка оказалась равна 7%.

3. На рис. 2.3 приведен пример добавления трехуровневого объекта к одноуровневому фону, достаточно близко отражающий случай появления дома с четырехскатной крышей. Параметры исходных изображений $S = 441$, $S_1 = 40$, $S_2 = 60$, $S_3 = 20$, $g_1^0 = 255$, $g_2^0 = 204$, $g_3^0 = 153$, $g_0^0 = 56$ (изображение f^0 считается однотонным). Оба изображения содержат дискретный шум, распределение которого аппроксимируется нормальным распределением с параметрами $N(0, 10)$. На рис. 2.6 изображены графики распределений $r^+(i)$ и $r^-(i)$. Количество пиков для распределения $r^+(i)$ равно двум, поскольку два пика слились в один (более высокий). Разделимость пиков, очевидно, связана с контрастном изображения: чем больше разница яркостей объекта и фона, тем дальше их пики друг от друга и, следовательно, выше качество работы алгоритмов. Более подробно этот вопрос обсуждается в п. 4.1.

Результаты поиска структурных различий для этих изображений приведены на рис. 2.7. Можно видеть, что появившийся дом найден точно, а мелкий шум легко отфильтровать по пло-

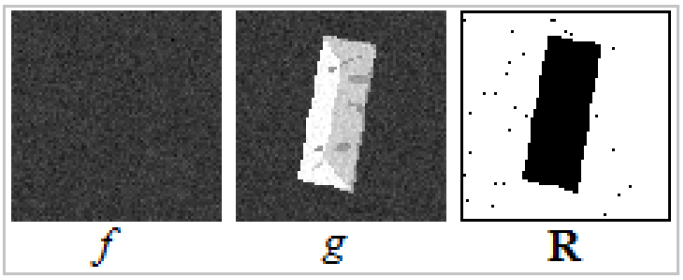

Рис. 2.7. Результат поиска структурных различий для изображений f и g, полученных путем добавления шума к однотонному изображению f^0 и содержащему трехуровневый объект изображению g^0 соответственно. Изображение $\mathbf{R}$ — это найденные с оптимальным порогом структурные различия (отмечены черным цветом).

щади. Параметр размера сканирующего окна d не использовался, поскольку все исходное изображение рассматривалось как локальное окно, а пороговое значение выбрано в соответствии с теоремой 2.4. Изображенные на рис. 2.7 найденные различия представляют собой наилучший из возможных результатов для данного алгоритма. Поскольку влияние шума на работу алгоритма достаточно существенно, то необходимо модифицировать алгоритм так, чтобы он подавлял случайные возмущения, что будет сделано в следующей главе.

2.7 Выводы по главе 2

В данной главе был предложен теоретический подход к формализации задачи поиска структурных различий изображений, опирающийся на теорию морфологического анализа изображений Ю. П. Пытьева. Полученные в рамках введенной модели формулы для распределений яркости зафиксированного пикселя разностного изображения позволяют получить точное значение оптимального порога алгоритма и оценки ошибок для него. Последнее весьма актуально для практических приложений: в реальном случае определить качество полученного результата возможно только с помощью ручной обработки снимков: построения разметки изменений и ее сравнения с выданными алгоритмом областями различий. Подобная операция достаточно трудозатратна, но ее можно избежать выбрав небольшой фрагмент и применив описанную выше методику построения распределений, которая даст оценку снизу для ошибок обнаружения.

Тем не менее, рассматриваемый в данной главе алгоритм поиска структурных различий является базовым, но не единственным. В следующей главе будет представлено подробное описание алгоритмов поиска структурных различий, основанных на различных функциях преобразования яркости.

Глава 3

Алгоритмы поиска структурных различий и их реализация

Данная глава посвящена описанию алгоритмов поиска структурных различий, относящихся к классу подходов, основанных на предварительном яркостном выравнивании исходных изображений для устранения несущественных изменений. Это выравнивание производится с помощью функций преобразования яркости. *Функция преобразования яркости* — это попиксельно действующий оператор относительной радиометрической коррекции исходных изображений, который строится для каждого положения сканирующего окна на основе наборов яркостей изображений в нем и применяется для "выравнивания" яркостей фрагментов с сохранением их структуры. Такой подход позволяет игнорировать не считающиеся структурными различиями объекты, изменившие свой цвет, но сохранившие форму.

В качестве функций преобразования яркости в данной работе используются морфологический проектор и его регуляризованный вариант, а также линейные и квадратичные функции. Как было показано во второй главе, задача поиска структурных различий сводится к построению изображений f' и g' по формулам обобщенной яркостной коррекции (2.1) и вычислению попиксельной разности преобразованных и исходных изображений. Так как характеристические функции ω_π, ω_1 и ω_2 принимают только значения ноль и бесконечность, то задача

$$f' = \arg\min(\|h - g\|^2 + \omega(h; f) \mid h \in \Phi)$$

эквивалентна задаче

$$f' = \arg\min(\|h - g\|^2 + \omega(h; f) \mid h\colon \omega(h; f) = 0),$$

т. е. можно говорить о минимизации по множеству $\{h\colon \omega(h; f) = 0\}$. Применение функций преобразования яркости изображения позволяет получить искомые изображения f' и g'.

Следует отметить, что в приведенных формулах в явном виде не участвуют яркости первого изображения f. Характеристические функции ω сравниваю структуры изображений как наборы точек, имеющих одинаковую яркость, безотносительно того, каково конкретное численное значение этой яркости. Подобный подход является идейным ядром яркостной коррекции, обеспечивающим независимость от соотношения средних яркостей входных изображений, отличающихся вследствие различных условий регистрации снимков.

Также в данной главе приводится альтернативный подход к яркостной коррекции изображений, основанный на глобальной оптимизации энергетической функции. Этот подход использует теорию марковских случайных полей для построения выровненных по яркости изображений и оперирует не с локальным окном, а с окрестностью каждого пикселя. Реализация основанного на этом подходе алгоритма поиска структурных различий выполнена с помощью метода имитации отжига [8]. Сравнение всех предложенных алгоритмов приведено в главе 4.

Однако использование яркостной коррекции изображений не всегда обязательно, что показывает метод, использующий вычисление "степени биективности". Данный алгоритм основан на определениях 2.2, 2.3 и обладает рядом специфических качеств, которые делают его недостаточно подходящим к непосредственному поиску структурных различий в общем случае, но крайне полезным в некоторых частных ситуациях: например, для геометрической коррекции исходных изображений или для поиска различий малой площади (в несколько пикселей).

В конце главы приводится асимптотическая оценка времени работы алгоритмов, описание программного комплекса, разработанного для поиска структурных различий на разновременных космических снимках, который был интегрирован в среду визуализации и обработки данных дистанционного зондирования Земли ENVI [38]. Также обсуждается применение высокопроизводительных графических процессоров для ускорения работы алгоритмов.

3.1 Общая схема алгоритма поиска структурных различий

Как было указано ранее, алгоритм поиска структурных различий должен находить изменения, связанные с появлением, исчезновением или изменением формы объектов. Для исключения влияния условий регистрации изображений и смены цвета объектов применяются специальные функции преобразования яркости. На вход алгоритм получает два геометрически выровненных снимка одинакового размера, как правило полутоновых, однако для морфологического проектора разработан вариант, работающий с цветными изображениями. Наконец, алгоритмы, основанные на различных функциях преобразования яркости, могут работать в группе, уточняя работу друг друга.

Общая схема алгоритма:

1. Исходные изображения сканируются с шагом 1 локальным окном заданного размера $d \times d$. Выбор подходящего размера окна подробно описан в главе 4. Центральную точку этого окна будем обозначать x_c. В случае необходимости производится дополнительное геометрическое выравнивание фрагментов (с помощью "степени биективности").

2. Для каждого положения окна по двум фрагментам сравниваемых изображений f и g строятся две функции преобразования яркости F_{fg} и F_{gf}. Необходимость использования схемы с двумя функциями связана с несимметричностью функций преобразования яркости и позволяет добиться симметризации результата, т. е. становится неважным, сравнивается первое изображение со вторым или наоборот.

3. С помощью функций F_{fg} и F_{gf} строятся преобразованные изображения $f' = F_{fg}(f)$ и $g' = F_{gf}(g)$, при этом яркость изображения f' "выровнена" по яркости изображения g с сохранением структуры изображения f, и аналогично — для изображения g'.

4. Строятся разностные изображения $R_{fg}(x) = |f'(x) - g(x)|$ и $R_{gf}(x) = |g'(x) - f(x)|$. Для этих изображений яркость точки характеризует величину структурного несоответствия исходных изображений, т. е. чем ярче точка, тем вероятнее, что в ней присутствует структурное различие.

5. Для завершения симметризации строится результирующее разностное изображение R, яркость каждой точки которого есть максимум яркостей точек разностных изображений с соответствующими координатами: $R(x) = \max(R_{fg}(x), R_{gf}(x))$.

6. Производится пороговая обработка изображения R, параметр порога будем обозначать T. Если стоит задача локализации структурного различия, то оценивается яркость центральной точки: из выполнения условия $R(x_c) \geq T$ следует, что в данной точке присутствует структурное различие, и в соответствующую точку результирующего выходного изображения $\mathbf{R}(x_c)$ записывается значение 1; если условие не выполняется, то записывается значение 0. Для задачи обнаружения структурного различия условие $R(x) \geq T$ проверяется для каждой точки сканирующего окна. Если число таких точек значительно, то считается, что в данном окне обнаружено структурное различие, и в $\mathbf{R}(x_c)$ записывается значение 1, иначе 0. Таким образом, обработка каждого сканирующего окна дает на результирующем изображении одну точку, яркость которой 1 или 0. Далее из этих точек формируются связные области структурных различий — конечный результат. В целях борьбы со случайными выбросами можно фильтровать найденные области, отбрасывая те, площадь которых незначительна.

На рисунке 3.1 представлена графическая схема, соответствующая описанию алгоритма поиска структурных различий.

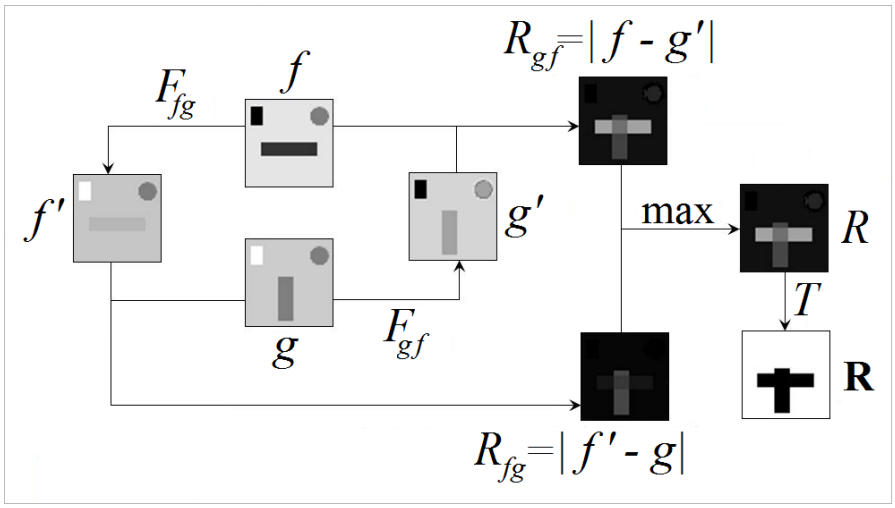

Рис. 3.1. Общая схема алгоритма поиска структурных различий изображений. Приведенные примеры изображений соответствуют работе алгоритма поиска различий на основе морфологического проектора.

Обратим внимание на несимметричность функции преобразования яркости. Рассмотрим следующий пример: пусть первое изображение f состоит из двух уровней яркости $L_f(i_1)$ и $L_f(i_2)$, а второе изображение g — из одного уровня яркости $L_g(i)$. В соответствии со введенной во второй главе формализацией, на изображении f присутствует структурное различие. Применим к этой паре изображений алгоритм, основанный, например, на морфологическом проекторе, который усредняет яркости второго изображения (в данном случае это однотонное изображение g) по уровням яркости первого. Таковых уровней два, и оба на втором изображении соответствуют одному и тому же уровню яркости g. Таким образом, полученное преобразованное изображение $f' = g$, а разностное изображение $R_{fg} \equiv 0$, т. е. никаких структурных различий обнаружить невозможно. Поэтому далее используется симметризация: сравнение не только первого изображения со вторым, но и наоборот. Например, на рис. 3.1 можно видеть, что присутствующий на изображении f горизонтально ориентированный черный прямоугольник сохраняется на изображении f', но исчезает на g'.

3.2 Алгоритм поиска структурных различий на основе морфологического проектора

В данном пункте рассматривается алгоритм поиска структурных различий, основанный на морфологическом проекторе, а также его модификации. Общая схема алгоритма описана в предыдущем пункте. В качестве функции преобразования яркости используется морфоло-

гический проектор (1.1), предложенный Ю. П. Пытьевым в работах [39, 40] и представляющий собой усреднение яркости второго изображений по уровням яркости первого. На рис. 3.2 приводится пример работы данного алгоритма для двух реальных полутоновых снимков городской застройки.

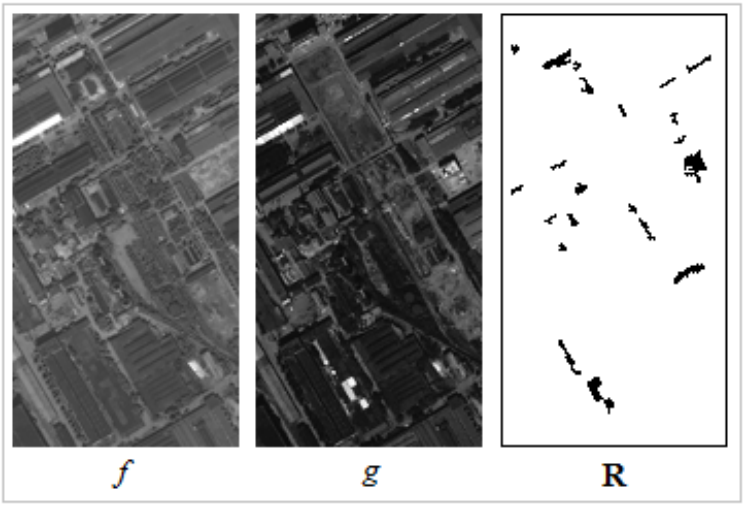

Рис. 3.2. Пример работы алгоритма поиска структурных различий на основе морфологического проектора: здесь f и g — исходные снимки, $\mathbf{R}$ — найденные структурные различия (отмечены черным цветом). Параметры алгоритма выбраны вручную: $d = 31$, $T = 45$.

Из представленного примера видно, что алгоритм показывает достаточно неплохую работу: пропусков и ложных тревог относительно немного. Однако также очевидно то, что алгоритм очень чувствителен к присутствующему на изображении шуму, поэтому в дальнейшем будет предложено несколько модификаций данного метода, обладающих повышенной устойчивостью к шуму. Особенностью данного алгоритма является то, что он не учитывает связность присутствующих на изображениях областей. Подобное поведение можно считать определенным недостатком алгоритма, для борьбы с котором существует два пути. Первый — это явное нахождение уровня яркости, соответствующее центральной точке сканирующего окна x_c: подобный подход рассмотрен в пункте 3.5. Второй путь заключается в использовании весовой функции, убывающей к границе окна: данный метод представлен в следующем пункте.

3.2.1 Регуляризация морфологического проектора

В данном пункте рассматривается вариант регуляризации морфологического проектора, предложенного Ю. П. Пытьевым, а также его применение в задаче поиска структурных разли-

чий. Еще раз приведем формулу (1.1) проектора:

$$P_f g(x) = \sum_i \frac{\sum\limits_{x' \in X} g(x') \cdot \chi_i^f(x')}{\sum\limits_{x' \in X} \chi_i^f(x')} \cdot \chi_i^f(x), \text{ где } \chi_i^f(x) = \begin{cases} 1, \text{ если } f(x) = i, \\ 0, \text{ иначе.} \end{cases}$$

Общая схема алгоритма описана в пункте 3.1. Подход, связанный с применением морфологического проектора, оптимален в случае отсутствия шума, т. е. когда каждый уровень яркости соответствует определенному объекту. Но при наличии на изображениях шума появляются "ложные" уровни яркости, которые морфологический проектор будет сохранять, что негативно сказывается на результате поиска структурных изменений.

Для устранения влияния шума требуется прибегать к методам регуляризации. В статье [6] описан один из вариантов борьбы с шумом, основанный на морфологической сегментации исходных изображений. В диссертации предлагается другой подход — регуляризация морфологического проектора, идея которой состоит в сглаживании характеристических функций уровней яркости.

Сначала выполняется сглаживание яркости второго изображения. Для зафиксированной яркости i усреднение яркостей второго изображения будет проводиться не по множеству точек, соответствующих данному уровню яркости на первом изображении, а по всему множеству пикселей (полю зрения X) с определенными весами. Вес определяется следующей формулой:

$$\tilde{\chi}_i^f(x) = \exp\left(-\frac{(f(x) - i)^2}{\sigma_c^2}\right),$$

где σ_c — параметр сглаживания. Таким образом, в точках, соответствующих зафиксированному уровню яркости i первого изображения, вес будет равен 1 (как и для морфологического проектора), а в остальных точках — некоторому числу от 0 до 1, в зависимости от того, насколько близка яркость текущей точки x к яркости i (для морфологического проектора в случае несовпадения яркостей было строго 0).

Для борьбы с граничными эффектами и получения более гладкого результата поиска структурных изменений в весовую формулу добавляется еще один множитель, уменьшающий вес точки в зависимости от ее расстояния до центра изображения (фактически, до центра сканирующего окна):

$$\tilde{\chi}_i^f(x) = \exp\left(-\frac{(f(x) - i)^2}{\sigma_c^2}\right) \cdot \exp\left(-\frac{(x - x_c)^2}{\sigma_d^2}\right), \tag{3.1}$$

где x_c —центральная точка сканирующего окна, σ_d — параметр сглаживания. Полученное выражение представляет собой билатеральный фильтр, разработанный для сглаживания изображений с сохранением границ [116].

Формула для проектора приобретает следующий вид:

$$P_f g(x) = \sum_i \frac{\sum\limits_{x' \in X} g(x') \cdot \tilde{\chi}_i^f(x')}{\sum\limits_{x' \in X} \tilde{\chi}_i^f(x')} \cdot \chi_i^f(x),$$

но здесь по-прежнему сохраняются искаженные шумом уровни яркости первого изображения. Поэтому для каждого уровня яркости i первого изображения строится изображение $\tilde{f}_i$, размеры которого совпадают с размерами исходных изображений, а яркости определяются по формуле:

$$\tilde{f}_i(x) = \frac{\sum\limits_{x' \in X} g(x') \cdot \tilde{\chi}_i^f(x')}{\sum\limits_{x' \in X} \tilde{\chi}_i^f(x')} \cdot \tilde{\chi}_i^f(x),$$

т. е. яркость точки x на изображении $\tilde{f}_i$ вычисляется по сглаженному изображению g и берется с определенным весом. Вес будет максимален, если на изображении f яркость этой точки равна i, и тем меньше, чем больше разность между яркостями $f(x)$ и i.

После этого все изображения $\tilde{f}_i$ попиксельно складываются, и яркость каждой точки делится на ее суммарный вес. Получается итоговая формула регуляризованного морфологического проектора

$$P_f g(x) = \frac{\sum\limits_i \tilde{f}_i(x)}{\sum\limits_j \tilde{\chi}_i^f(x)} = \sum_i \left(\frac{\sum\limits_{x' \in X} g(x') \cdot \tilde{\chi}_i^f(x')}{\sum\limits_{x' \in X} \tilde{\chi}_i^f(x')} \cdot \frac{\tilde{\chi}_i^f(x)}{\sum\limits_j \tilde{\chi}_i^f(x)} \right), \tag{3.2}$$

где $\tilde{\chi}_i^f$ вычисляется по формуле (3.1).

Таким образом, полученный проектор сохраняет преимущества морфологического проектора (устранение влияния условий регистрации) и при этом подавляет случайный шум, присутствующий на исходных изображениях. На рис. 3.3 представлены результаты работы алгоритма. Снимки крайние слева — пара исходных изображений f и g городской застройки, сделанные с интервалом 4 года. В центре — те же изображения, преобразованные с помощью регуляризованного морфологического проектора. Далее взяты разность изображений f с $P_g f$ и g с $P_f g$. Можно видеть, что для этих пар изображений влияние условий регистрации устранено и их оставшееся несоответствие и есть структурные различия. Справа вверху — найденные алгоритмом области структурных различий (в белой рамке — увеличенный фрагмент).

Для сравнения на рис. 3.3 справа внизу приведены области структурных различий для тех же исходных изображений, выделенные алгоритмом на основе морфологического проектора. Оба метода нашли одни и те же области интереса, но регуляризация проектора позволила сделать эти области более однородными по яркости, без дырок и с сохранением точности границ.

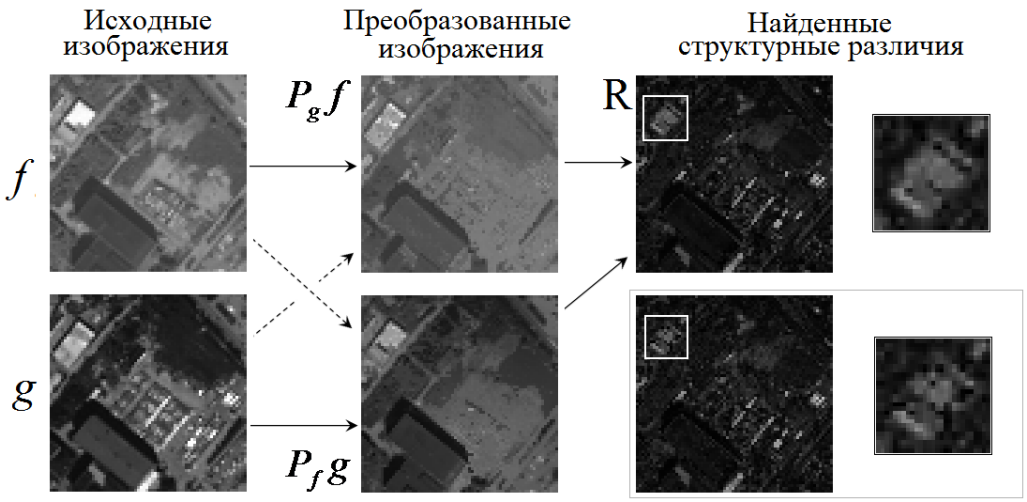

Рис. 3.3. Результаты работы алгоритма: исходные изображения, преобразованные изображения и разностные изображения для алгоритмов на основе морфологического проектора и его регуляризованного варианта. В белой рамке —увеличенный фрагмент.

Далее было проведено сравнение предлагаемого варианта регуляризации с обычным сглаживанием исходных изображений. Для этого исходные изображения сглаживались окном размера 3×3 с коэффициентами, вычисленными по формуле Гаусса

$$\Gamma(x, y) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right).$$

По результирующему выходному изображению $\mathbf{R}$, содержащему найденные для сглаженных изображений области структурных различий, были построены соответствующие ROC-кривые по порогу и подсчитаны площади под этими ROC-кривыми (более подробно об анализе ROC-кривых будет рассказано в четвертой главе): для морфологического проектора без сглаживания исходных изображений площадь равна 0.8473; со сглаживанием — 0.8662; для регуляризованного проектора и исходных изображений — 0.8522; для сглаженных изображений — 0.8711.

Можно сделать следующие выводы: положительные свойства предложенного метода (по сравнению с морфологическим проектором) заключаются в том, что он дает более “гладкий” результат — более однородные, сглаженные области найденных структурных различий без потери точности границ областей изменений. Также предложенный метод лучше с точки зрения анализа ROC-кривых. Отрицательные стороны — значительное время работы и необходимость настройки дополнительных параметров. Более полное сравнение методов и оценка оптимальных параметров работы приведены в главе 4.

3.2.2 Морфологический проектор для цветных изображений

Цель данной работы — решение задачи поиска структурных различий для полутоновых снимков земной поверхности. Тем не менее, определенный интерес представляет решение данной задачи для цветных изображений. Простейшим вариантом является поканальная обработка цветных снимков предложенными методами. Однако, такой подход не дает ничего нового по сравнению со схемой обесцвечивания и применения стандартных методов, в то время как цветное изображение содержит гораздо больше полезной информации. Данный вопрос несколько шире темы диссертации, и в данном пункте будет представлено лишь небольшое ее рассмотрение, связанное с предложением варианта морфологического проектора для цветных изображений.

Идея предлагаемого метода идентична вышеописанным подходам: первоначально требуется произвести радиометрическую коррекцию входных изображений. Ю. П. Пытьевым в работе [40] тема устранения влияния условий регистрации для цветных изображений раскрыта достаточно широко. Здесь будет предложен один из возможных вариантов проектора.

Поскольку теперь речь будет идти о цветных изображениях, то придется внести изменения во введенное в пункте 1.1 определение изображения. А именно, в данном пункте под *изображением* $\vec{f}$ будем понимать функцию $\vec{f}\colon X \to \mathbb{R}^3$, а значение $\vec{f}(x) = (f_R(x), f_G(x), f_B(x))$ будем называть *яркостью* изображения $\vec{f}$ в точке $x \in X$, где f_R, f_G и f_B — яркости по красному, зеленому и синему каналу, соответственно.

Суть морфологического проектора заключается в усреднении яркостей второго изображения по уровням яркости первого (1.1). Однако представляется довольно затруднительным задать уровень яркости как некоторое множество в пространстве $\mathbb{R}^3$: любое разбиение является и подходящим, и неудачным, что делает задачу такого разбиения неоднозначной. Однако известно, что предлагаемый в данной работе алгоритм оценивает центральную точку окна, поэтому можно задавать уровень яркости как точки, близкие по яркости к центральной x_c:

$$\hat{L}_{\vec{f}}(x_c) = \{x \in X \mid \max(|f_R(x) - f_R(x_c)|, |f_G(x) - f_G(x_c)|, |f_B(x) - f_B(x_c)|) \le t_c\},$$

где t_c — параметр яркостной близости. В таком случае каждой точке первого изображения соответствует свой набор точек и свой уровень яркости, но эти наборы точек могут иметь непустое пересечение. Тогда формула морфологического проектора для цветных изображений выглядит следующим образом:

$$P_{\vec{f}}\vec{g}(x) = \sum_{\vec{i}} \frac{\sum\limits_{x' \in X} \vec{g}(x') \cdot \hat{\chi}^{\vec{f}}(x', x)}{\sum\limits_{x' \in X} \hat{\chi}^{\vec{f}}(x', x)} \cdot \chi_i^{\vec{f}}(x), \tag{3.3}$$

где $\vec{i} = (i_R, i_G, i_B)$, $\quad \chi_i^{\vec{f}}(x) = \begin{cases} 1, & \text{если } \vec{f}(x) = \vec{i}, \\ 0, & \text{иначе}, \end{cases} \qquad \hat{\chi}^{\vec{f}}(x', x) = \begin{cases} 1, & \text{если } x' \in \hat{L}_{\vec{f}}(x), \\ 0, & \text{иначе}. \end{cases}$

Алгоритм поиска структурных различий с функцией преобразования яркости (3.3) полностью совпадает с описанным в пункте 3.1. На рис. 3.4 представлены результаты его работы для двух цветных изображений городской застройки.

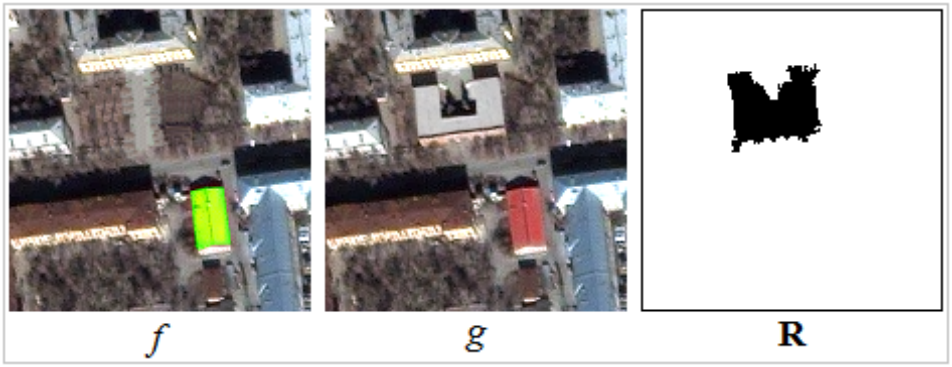

Рис. 3.4. Результаты работы алгоритма поиска структурных различий для цветных изображений: исходные изображения f и g и найденные области структурных различий $\mathbf{R}$ (отмечены черным цветом).

Параметры алгоритма были выбраны вручную: яркостная близость для линии уровня $t_c = 16$, порог $T = 55$. Можно видеть, что алгоритм успешно нашел появившийся дом, проигнорировав строение, крыша которого изменила свой цвет. Таким образом, можно говорить о применимости данного метода для поиска структурных различий на разновременных цветных снимках земной поверхности. Для улучшения качества его работы можно провести регуляризацию проектора, как в пункте 3.2.1.

3.3 Алгоритмы поиска структурных различий для отдельных классов функций яркостной коррекции изображений

Как было указано выше, морфологический проектор полезен для устранения влияния условий регистрации изображений, но обладает недостатком — высокой чувствительностью к шуму, поскольку весь имеющийся сигнал считается полезным, и потому стремится его сохранить. Если на первом изображении присутствует слишком много уровней яркости, результат его преобразования будет очень близок ко второму изображению как по яркости, так и по структуре. Следовательно, на такой паре изображений не удастся обнаружить структурные различия, даже если они есть. Например, для двух изображений размером 16×16 пикселей, принимающих все 256 значений яркости, после преобразования с помощью морфологического

проектора на изображениях f' и g' значения яркостей пикселей будут совпадать со значениями соответствующих пикселей g и f. Поэтому никаких структурных различий на этих изображениях обнаружено не будет. В то же время, как показано на рис. 3.5, можно так подобрать эти изображения, что на них будет присутствовать заметное структурное различие: здесь (а) и (б) — способ раскраски, заключающийся в заполнении пикселей по диагонали яркостями с шагом 1, (в) и (г) — получившиеся изображения. В реальном случае аналогичная ситуация возможна, если на изображениях присутствует сильный шум.

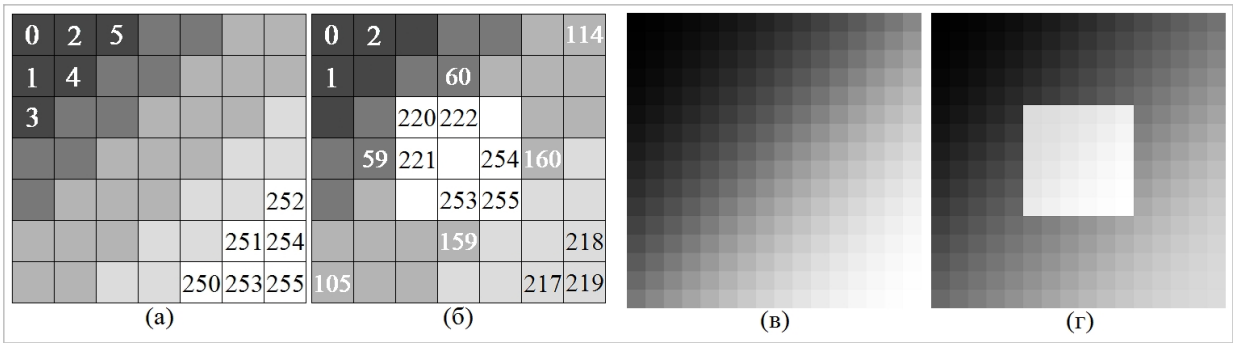

Рис. 3.5. Пример изображений, на которых морфологический проектор работает плохо. Здесь (а) и (б) — способ построения изображений, (в) и (г) — изображения, полученные по схемам (а) и (б), соответственно.

Один из основных вариантов борьбы с этим недостатком морфологического проектора — сглаживание исходных изображений. Однако такой подход связан с потерей информации и размытием границ, что негативно сказывается на качестве локализации структурных различий. Второй путь, предложенный в пункте 3.2.1, — регуляризация морфологического проектора, позволяющая успешно подавлять случайный шум без потери четкости границ и, следовательно, качества работы алгоритма. Ценой, которую приходится платить за качественную работу алгоритма, основанного на регуляризованном варианте морфологического проектора, является сложность настройки и длительное время работы.

Поэтому предлагается еще один подход, связанный с ограничением класса допустимых преобразований $\mathfrak{F}$ линейными и квадратичными функциями. Суть данного предложения такова: морфологический проектор является оптимальным приведением изображений друг к другу, что доказано в [39, 40], но эта оптимальность будет лишь тогда, когда класс допустимых преобразований $\mathfrak{F}$ ограничен лишь классом борелевских функций. В то же время, следствием оптимальности проектора является точная подгонка изображений друг к другу с сохранением всех уровней яркости. Соответственно, если ограничить класс преобразований, например, линейными функциями, то проектор уже не будет оптимальным из-за потери незначительных данных,

скорее всего являющихся несущественным случайным шумом, и изображения будут приведены друг к другу более “грубо”. Как указано в обзоре литературы (п. 1.2), линейная функция преобразования яркости достаточно часто и небезуспешно применяется различными авторами для яркостной коррекции исходных снимков [54, 112, 121]. В данной диссертации ее применение получает конкретное объяснение, а также для сравнения исследуется класс квадратичных функций.

Описание алгоритмов, основанных на линейных и квадратичных функциях преобразования яркости, соответствует схеме из пункта 3.1, а построение функций приводится в следующих двух пунктах. На рис. 3.6 представлено сравнение результатов их работы с результатами работы алгоритма, основанного на регуляризованном морфологическом проекторе для изображений, построенных по схеме, приведенной на рис. 3.5.

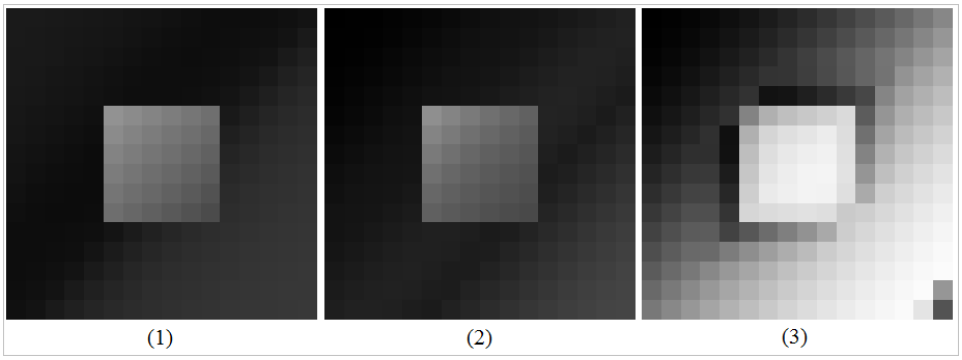

Рис. 3.6. Разностные изображения, полученные для изображений, представленных на рис. 2.10 для алгоритмов поиска структурных различий, основанных на: (1) — линейной функции преобразования яркости, (2) — квадратичной функции преобразования яркости, (3) — регуляризованном морфологическом проекторе.

Результирующее изображение для алгоритма, основанного на морфологическом проекторе, представляет собой тождественный ноль, т. е. никаких структурных различий не обнаруживается. На разностных изображениях, соответствующих линейной и квадратичной функциям преобразования яркости (рис. 3.6 (1) и 2.11 (2)), объект интереса выделен достаточно явно, и его легко локализовать с помощью пороговой обработки (при этом линейная функция дала чуть лучший результат). А вот результат работы алгоритма с регуляризованным морфологическим проектором (рис. 3.6 (3)) показал результаты неоднозначные: с одной стороны, объект интереса выделен четко, с другой — в правом нижнем углу изображения присутствует ложное срабатывание. Следовательно, сглаживание (изображений или функций, лежащих в основе алгоритмов) не всегда дает положительный результат, поэтому простые алгоритмы, основанные на определенных классах функций преобразования яркости, также имеют право на жизнь. Об-

щее экспериментальное сравнение качества и времени работы всех алгоритмов приводится в этой и четвертой главах диссертации.

3.3.1 Класс линейных функций преобразования яркости

Использование линейных функций для яркостной коррекции изображений — один из базовых подходов, применяющихся в задаче поиска структурных различий, причем вариантов его применения достаточно много. Различие между ними заключается в способе оценки коэффициентов линейной функции. В данной работе коэффициенты находятся с помощью метода наименьших квадратов по двум фрагментам изображений, ограниченных сканирующим окном:

$$\sum_{x=1}^{S_d}(g(x) - (k \cdot f(x) + b))^2 \to \min_{k,b},$$

где k и b — коэффициенты линейного преобразования. Решая данную задачу, получаем известные соотношения:

$$k = \frac{S_d \sum\limits_{x=1}^{S_d} f(x)g(x) - \sum\limits_{x=1}^{S_d} f(x) \sum\limits_{x=1}^{S_d} g(x)}{S_d \sum\limits_{x=1}^{S_d} f^2(x) - \left(\sum\limits_{x=1}^{S_d} f(x)\right)^2}, \qquad b = \frac{\sum\limits_{x=1}^{S_d} g(x) - k \sum\limits_{x=1}^{S_d} f(x)}{S_d}.$$

Здесь $S_d = d \times d$ — площадь сканирующего окна, а под f и g понимаются ограниченные локальным окном фрагменты, а не целые изображения.

С найденными коэффициентами линейная функция преобразования яркости имеет вид:

$$f'(x) = k \cdot f(x) + b. \tag{3.4}$$

По аналогии с теоремой 2.1 для морфологического проектора, следующее утверждение показывает, что использование линейной функции преобразования яркости (3.4) эквивалентно применению формул обобщенной яркостной коррекции (2.1).

Утверждение 3.1. 1. *Использование для построения выровненного по яркости изображения линейной функции преобразования яркости, коэффициенты которой находятся с помощью метода наименьших квадратов, эквивалентно*

$$f' = \arg\min(\|h - g\|^2 + \omega_1(h; f) \mid h \in \Phi),$$

где $\omega_1(h; f) = 0$, *если* $h(x) = kf(x) + b$ *для некоторых вещественных* k, b, *и иначе равна бесконечности.*

Д о к а з а т е л ь с т в о. В соответствии с формулой (3.4), для построения выровненного по яркости изображения используется линейная функция преобразования яркости, коэффициенты которой находятся с помощью метода наименьших квадратов:

$$\sum_{x=1}^{S_d}((k \cdot f(x) + b) - g(x))^2 \to \min_{k,b},$$

т. е.

$$\sum_{x=1}^{S_d}((k' \cdot f(x) + b') - g(x))^2 = \min_{k,b}\left(\sum_{x=1}^{S_d}((k \cdot f(x) + b) - g(x))^2\right),$$

или, обозначив $f'(x) \equiv k' \cdot f(x) + b'$ и $h_{k,b}(x) \equiv k \cdot f(x) + b$,

$$\|f'(x) - g(x)\|^2 = \min_{k,b}(\|h_{k,b}(x) - g(x)\|^2).$$

Таким образом, минимум ищется среди изображений, полученных из изображения f путем применения всех возможных линейных преобразования яркости, т. е.

$$f'(x) = \arg\min(\|h(x) - g(x)\|^2 \mid h(x) = k \cdot f(x) + b; k, b \in \mathbb{R}).$$

А это совпадает с условием равенства нулю характеристической функции $\omega_1(h, f)$. Тогда $h(x) = k \cdot f(x) + b$ можно переписать как $\{h \mid h$ — любое, но $\omega_1(h; f) = 0\}$. Следовательно, изображение f' будет вычисляться по следующей формуле:

$$f' = \arg\min(\|h - g\|^2 \mid h \text{ — любое, но } \omega_1(h; f) = 0),$$

или, поскольку ω_1 принимает только два значения (ноль, бесконечность),

$$f' = \arg\min(\|h - g\|^2 + \omega_1(h; f) \mid h \text{ — любое}). \blacksquare$$

На рис. 3.7 приведен результат работы алгоритма поиска структурных различий на основе линейной функции преобразования яркости для двух реальных разновременных снимков городской застройки. Можно видеть, что алгоритм достаточно точно нашел имеющиеся структурные различия, справившись с присутствующим случайным шумом.

3.3.2 Класс квадратичных функций преобразования яркости

Квадратичная функция преобразования яркости строится аналогично функции линейной. Коэффициенты находятся также с помощью метода наименьших квадратов:

$$\sum_{x=1}^{S_d}(g(x) - (k \cdot f^2(x) + b \cdot f(x) + c))^2 \to \min_{k,b,c},$$

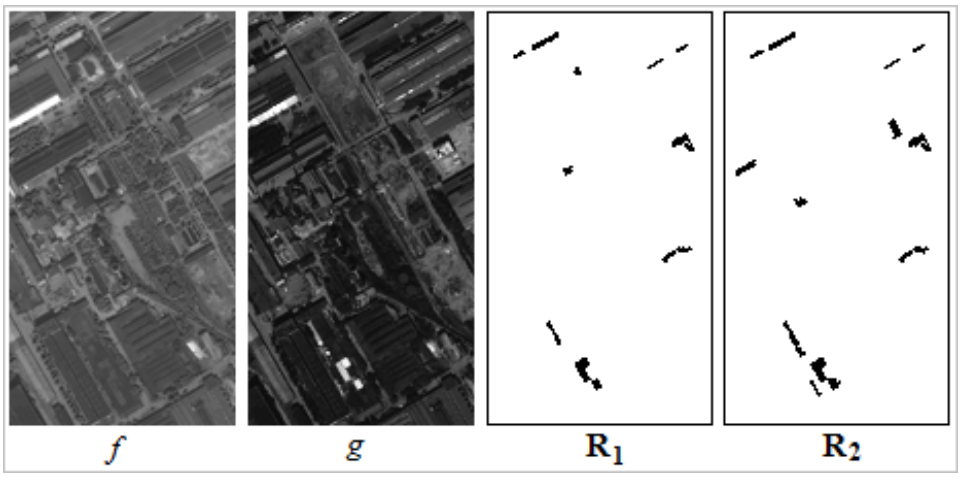

Рис. 3.7. Пример работы алгоритма поиска структурных различий на основе линейной ($\mathbf{R_1}$) и квадратичной ($\mathbf{R_2}$) функций преобразования яркости: здесь f и g — исходные снимки, $\mathbf{R_1}$ и $\mathbf{R_2}$ — найденные структурные различия (отмечены черным цветом). Параметры алгоритмов выбраны вручную: $d = 31$, $T = 50$.

где k, b и c — коэффициенты квадратичного преобразования. Решая данную задачу, получаем:

$$k_1 = \left(S_d \sum_{x=1}^{S_d} f^2(x) - \sum_{x=1}^{S_d} f(x)\right) \cdot \left(S_d \sum_{x=1}^{S_d} f^2(x)g(x) - \sum_{x=1}^{S_d} f^2(x) \sum_{x=1}^{S_d} g(x)\right)$$

$$- \left(S_d \sum_{x=1}^{S_d} f(x)g(x) - \sum_{x=1}^{S_d} g(x)\right) \cdot \left(S_d \sum_{x=1}^{S_d} f^3(x) - \sum_{x=1}^{S_d} f(x) \sum_{x=1}^{S_d} f^2(x)\right),$$

$$k_2 = \left(S_d \sum_{x=1}^{S_d} f^2(x) - \sum_{x=1}^{S_d} f(x)\right) \cdot \left(S_d \sum_{x=1}^{S_d} f^4(x) - \left(\sum_{x=1}^{S_d} f^2(x)\right)^2\right)$$

$$- \left(S_d \sum_{x=1}^{S_d} f^3(x) - \sum_{x=1}^{S_d} f^2(x)\right) \cdot \left(S_d \sum_{x=1}^{S_d} f^3(x) - \sum_{x=1}^{S_d} f(x) \sum_{x=1}^{S_d} f^2(x)\right),$$

$$k = \frac{k_1}{k_2}, \qquad b = \frac{S_d \sum\limits_{x=1}^{S_d} f(x)g(x) - \sum\limits_{x=1}^{S_d} g(x) - k\left(S_d \sum\limits_{x=1}^{S_d} f^3(x) - \sum\limits_{x=1}^{S_d} f^2(x)\right)}{S_d \sum\limits_{x=1}^{S_d} f^2(x) - \sum\limits_{x=1}^{S_d} f(x)},$$

$$c = \frac{\sum\limits_{x=1}^{S_d} g(x) - k \sum\limits_{x=1}^{S_d} f^2(x) - b \sum\limits_{x=1}^{S_d} f(x)}{S_d}.$$

Здесь $S_d = d \times d$ — площадь сканирующего окна, а под f и g также понимаются ограниченные локальным окном фрагменты, а не целые изображения.

С найденными коэффициентами квадратичная функция преобразования яркости имеет вид:

$$f'(x) = k \cdot f^2(x) + b \cdot f(x) + c. \tag{3.5}$$

Также, как и для класса линейных функций, можно показать, что использование квадратичной функции преобразования яркости (3.5) эквивалентно применению формул (2.1).

Утверждение 3.2. 1. *Использование для построения выровненного по яркости изображения квадратичной функции преобразования яркости, коэффициенты которой находятся с помощью метода наименьших квадратов, эквивалентно*

$$f' = \arg\min(\|h - g\|^2 + \omega_2(h; f) \mid h \in \Phi),$$

где $\omega_2(h; f) = 0$*, если* $h(x) = kf^2(x) + bf(x) + c$ *для некоторых вещественных* k*,* b*,* c*, и иначе равна бесконечности.*

Доказательство аналогично доказательству утверждения 3.1. ■

На рис. 3.7 приведен результат работы алгоритма поиска структурных различий на основе квадратичной функции преобразования яркости. Результаты работы несколько отличаются от примера с линейной функцией: найдено больше объектов, что свидетельствует о том, что квадратичная функция преобразования яркости сохраняет больше информации, чем линейная.

В заключение можно сказать, что использование конкретных классов функций преобразования яркости позволяет успешно бороться даже с самым значительным случайным шумом, зачастую превосходя в этом даже сглаживание / регуляризацию. Дальнейшее исследование и сравнение алгоритмов приводится в главе 4.

3.4 Алгоритм поиска структурных различий на основе глобальной оптимизации

В предыдущих частях работы описаны алгоритмы, чьи структурные характеристические функции ω принимают конечное число значений. Основанные на них методы осуществляют приведение по яркости изображений друг к другу, и в случае морфологического проектора такое приведение является оптимальным. Однако строгое сохранение всех уровней яркости может оказать негативное влияние на результат работы алгоритма при наличии на входных изображениях случайного шума. Обойти этот недостаток можно либо регуляризацией проектора, либо ограничением класса допустимых преобразований яркости. Но и в этом случае алгоритм стремится сохранить структуру первого изображения: морфологический проектор — полностью,

остальные методы — частично. В таком случае возникает интерес создания алгоритма, в котором можно было бы выбирать соотношение между сохранением структуры первого изображения и точностью яркостного соответствия со вторым фрагментом. В данном пункте речь пойдет об использовании структурной характеристической функции ω_m, принимающей значения из некоторого отрезка, использование которой дает возможность варьировать то, на сколько близко изображения будут приведены друг к другу.

В основе алгоритма поиска структурных различий с непрерывной структурной характеристической функцией ω_m лежит, как и в предыдущих случаях, яркостная коррекция входных снимков, заключающаяся в приведении набора яркостей первого изображения к яркости второго с сохранением структуры первого изображения. Для улучшения результата также производится приведение каждого изображения к самому себе, что позволяет избавиться от присутствующего шума. Реализация такого алгоритма сводится к задаче глобальной минимизации функционалов (2.1), что эквивалентно минимизации следующей энергетической функции:

$$E(f'; f, g) = \|f'(x) - g(x)\|^2 + \omega_m(f'; f) =$$

$$= \sum_{x \in X} (f'(x) - g(x))^2 + \lambda \sum_{x \in X} \sum_{y \in N_x} G_f(f'(x) - f'(y), f(x) - f(y)), \tag{3.6}$$

$$G_f(i, j) = \begin{cases} \min(i^2, \alpha^2), & \text{если } |j| < \alpha, \\ 0, & \text{иначе,} \end{cases}$$

где N_x — окрестность точки x, λ и α — положительные параметры. Параметр λ отвечает за близость приведения изображения f к g: чем он больше, тем больше алгоритм будет стремиться сохранить структуру изображения f, из-за чего яркостное приведение f к g будет менее точным. Если же $\lambda \to 0$, то структура изображения f будет разрушаться для наиболее полного совпадения яркостей f и g (очевидно, что если $\lambda = 0$, то $f' = g$). Параметр α характеризует структуру изображения f, задавая допустимый перепад яркости между различными линиями уровня: если яркости двух точек отличаются меньше, чем на α, то они принадлежат одному уровню.

В функции E первое слагаемое отвечает за то, насколько близки два изображения по яркости. Второе слагаемое — это структурная характеристическая функция, играющая роль функции штрафа за нарушение структуры. Если на изображении f, к которому мы приводим изображение f' по структуре, яркости центральной точки x_1 и точки x_2 из ее окрестности близки (меньше α), то накладывается штраф, равный квадрату подобной разности в соответствующих точках f, но значение которого не больше квадрата α. В противном случае, т. е. это край, штраф не накладывается. Не имеет значения факт пропажи края, так как ищутся только появившиеся объекты, а поиск “исчезнувших” объектов осуществляется при симметризации

результата. Аналогичным образом осуществляется приведение второго изображения к первому, а также каждого из изображений к самому себе.

Для нахождения минимума функции $E(f'; f, g)$ будет использоваться аппарат марковских случайных полей (МСП) [83]. Одним из наиболее простых в реализации методов приближенного решения подобной задачи является алгоритм Метрополиса [8]. В качестве энергетической функции для данного метода будем использовать введенную функцию $E(f'; f, g)$. Общая схема алгоритма поиска структурных различий, основанного на глобальной оптимизации, существенно отличается от приведенной в пункте 3.1 и состоит из следующих шагов.

1. Моделируется начальная (пробная) конфигурация f'. Здесь конфигурация трактуется в терминах МСП — как реализация случайного поля. Фактически f' — это изображение с преобразованными яркостями, которое будет модифицироваться в процессе оптимизации энергетической функции. После проведения экспериментов было получено, что наилучшим первым приближением будет $f' = g$ при приведении по яркости изображения f к g, и $g' = f$ — при симметризации результата.

2. Производится корректировка конфигурации f' путем случайного изменения яркостей точек, что позволяет "выходить" из точек локального минимума. Пусть новая конфигурация обозначается f''. Если $E(f''; f, g) \leq E(f'; f, g)$, то f'' принимается в качестве новой конфигурации. Если $E(f''; f, g) > E(f'; f, g)$, то f'' принимается в качестве новой конфигурации с вероятностью $\exp((E(f'; f, g) - E(f''; f, g))/Temp)$, где $Temp$ — параметр температуры в схеме имитации отжига [8]. Если f'' отклоняется, то остается прежняя конфигурация f'.

3. Корректировка конфигурации производится достаточное количество раз (порядка 5000, как правило, достаточно; в [8] приведено доказательство сходимости данного метода). При этом на каждом шаге производится понижение параметра температуры $Temp$. "Охлаждение" в схеме имитации отжига осуществляется по следующему эвристическому правилу $Temp^{(t)} = 0.999 \cdot Temp^{(t-1)}$, где t — номер итерации ([8], глава 10).

4. Аналогичным образом строится изображение g', что обеспечивает симметризацию результата. Для устранения влияния шума также строятся изображения $\bar{f}'$ и $\bar{g}'$, являющиеся результатом оптимизации энергетических функций $E(\bar{f}'; f, f)$ и $E(\bar{g}'; g, g)$, соответственно.

5. Далее вычисляются разностные изображения $R_{fg}(x) = |f'(x) - \bar{g}'(x)|$ и $R_{gf}(x) = |g'(x) - \bar{f}'(x)|$. Для этих изображений яркость точки характеризует величину структурного несоответствия исходных изображений, т. е. чем ярче точка, тем вероятнее, что в ней присутствует структурное различие.

6. Для завершения симметризации строится результирующее разностное изображение R, яркость каждой точки которого есть максимум яркостей точек разностных изображений с соот-

ветствующими координатами: $R(x) = \max(R_{fg}(x), R_{gf}(x))$. Далее производится пороговая обработка изображения R. Если стоит задача локализации структурного различия, то оценивается яркость центральной точки: при выполнении условия $R(x_c) \geq T$ следует, что в данной точке присутствует структурное различие, и в соответствующую точку результирующего выходного изображения $\mathbf{R}(x_c)$ записывается значение 255; если условие не выполняется, записывается 0.

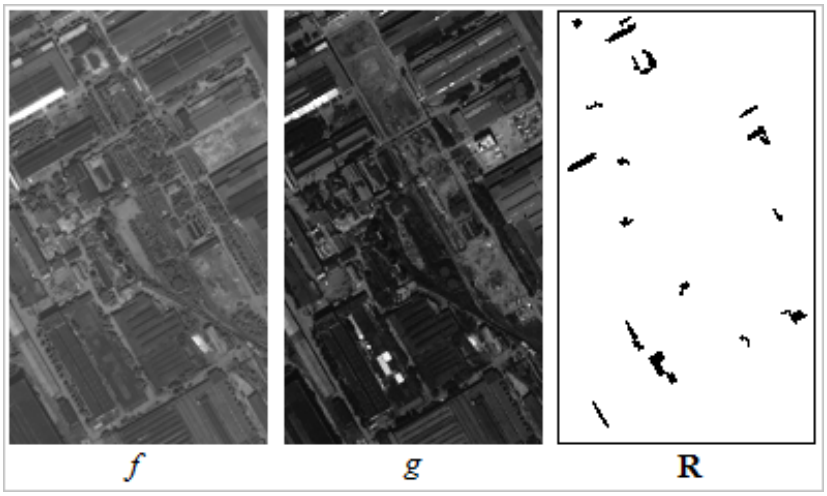

Рис. 3.8. Пример работы алгоритма поиска структурных различий на основе глобальной оптимизации: здесь f и g — исходные снимки, $\mathbf{R}$ — найденные структурные различия (отмечены черным цветом). Параметры алгоритма выбраны вручную: $d = 5$, $\lambda = 1000$, $\alpha = 15$ и $T = 50$.

На рис. 3.8 приведены результаты работы данного алгоритма. Параметр d теперь означает не размер сканирующего окна, а размер окрестности N_x точки x.

Достоинством данного алгоритма в сравнении с описанными ранее методами является большая чувствительность к изменениям в составе объектов. Однако, следствием этого становится и большая подверженность влиянию шума. Еще одним минусом можно считать использование наибольшего среди всех алгоритмов количества параметров.

3.5 Алгоритм поиска структурных различий без использования яркостной коррекции изображений

В соответствии с определением 2.2 структура изображения — это набор его множеств уровня; структурное различие согласно определению 2.3 — это набор точек, в которых яркости пары изображений, выровненных геометрически и по яркости, существенно различаются.

В целях повышения точности локализации найденных различий большая часть алгоритмов основана на работе со сканирующим окном и последующей оценкой степени различия для его центральной точки x_c. При доказательстве теорем в пункте 2.6 было отмечено, что при таком подходе к поиску различий на первом изображении можно рассматривать лишь уровень яркости, которому принадлежит центральная точка. Основываясь на этих рассуждениях, можно предложить подход, анализирующий на каждом из изображений яркости пикселей сканирующего окна относительно яркости центральной точки x_c, благодаря чему становится возможным не использовать яркостную коррекцию снимков. Такой алгоритм получил название “степени биективности”.

В основе вычисления “степени биективности” SB лежит следующая идея: при отсутствии структурных различий на изображениях малая окрестность яркости каждого пикселя изображения f должна перейти в малую окрестность яркости соответствующего пикселя изображения g, а точки вне этой окрестности не должны попасть в эту окрестность. Другими словами, при отсутствии структурных различий пиксели, на первом изображении близкие по яркости к $f(x_c)$ (попадающие в некоторую окрестность яркости точки x_c), на втором должны оказаться близки по яркости к $g(x_c)$. Необходимость задания окрестности связана с задачей подавления шума на изображениях. Формально это можно определить с помощью понятия штрафа. Для этого задаются два радиуса окрестностей — δ для первого и ε для второго изображения. Пиксели, яркость которых лежит за пределами (в) δ-окрестности границы яркости центральной точки x_c первого изображения, а на втором изображении (не) попадают в ε-окрестность яркости центральной точки, штрафуются на величину α (β) (рис. 3.9). Итоговая оценка в данной точке — “*степень биективности*“ — есть сумма штрафов для всех пикселей локального окна.

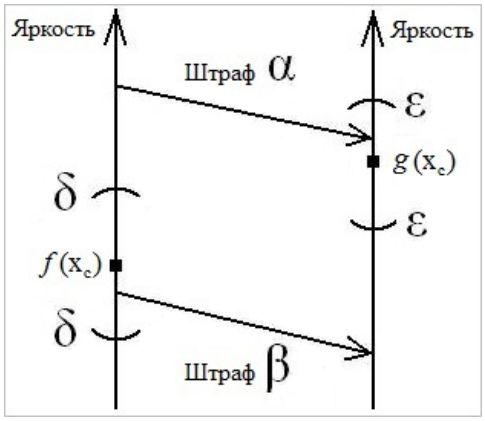

Рис. 3.9. Принцип вычисления “степени биективности” для двух изображений.

Дадим более строгое определение “степени биективности”. Пусть уровни яркости опреде-

ляются соотношениями:

$$L_f^0(x_c) = \{x \in X \mid |f(x) - f(x_c)| < \delta\},$$

$$L_g^0(x_c) = \{x \in X \mid |g(x) - g(x_c)| < \varepsilon\},$$

где X — множество пикселей сканирующего окна. Зададим функцию штрафа $h\colon X \to \mathbb{R}$. В частном случае каждая оказавшаяся за пределами заданной окрестности точка штрафуется на 1 ($h(\cdot) \equiv 1$), но можно использовать и другие весовые функции, например, функцию Гаусса, убывающую к границе окна. Тогда "степень биективности" вычисляется по формуле:

$$SB(x_c) = \sum_{x \in L^0} h(x), \tag{3.7}$$

где L^0 — симметрическая разность двух множеств: $L^0 = L_f^0(x_c) \triangle L_g^0(x_c)$.

В пункте 2.1 уже упоминалось, что попытки определения структурных различий на основе симметрической разности приводят к неоднозначности трактовок (рис. 2.1). Однако "степень биективности" вовсе не обязательно рассматривать как подход для поиска различий в самом общем случае. Преимущества данного метода проявляются в следующих случаях.

1. Поиск структурных различий, размеры которых составляют несколько пикселей. При анализе космосъемки такие выбросы почти наверняка являются шумом, и все вышеперечисленные алгоритмы будут стремиться их подавить. Однако, возможно появление практической задачи, где ключевую роль будут играть именно такие незначительные по площади различия, и вычисление "степени биективности" позволит их точно локализовать.

2. Локальная геометрическая коррекция исходных снимков. Качество привязки снимков друг к другу имеет критическую важность для правильности работы алгоритмов поиска различий. К сожалению, ситуации, когда на вход поступают снимки с несовершенной привязкой, вполне возможны, и потому необходимо учитывать этот фактор в алгоритме. Предлагается следующая стратегия использования "степени биективности": алгоритмы работают со сканирующим окном, размеры которого относительно невелики, и можно считать, что присутствующие искажения проявляются на размерах больше, чем величина стороны локального окна d. Тогда, прежде чем строить функции преобразования яркости, можно вычислить "степень биективности" для текущего положения окна и для всех окон со сдвигом по небольшой окрестности, например, 5×5. Определить эту величину лучше вручную, проанализировав снимок. Далее выбирается то положение окна, где величина SB минимальна, и дальнейшие расчеты проводятся для него, но результат обработки записывается по прежнему в центр окна в исходном положении.

На рис. 3.10 представлено сравнение результатов работы алгоритма поиска структурных различий изображений на основе морфологического проектора при отсутствии локальной

геометрической коррекции и в ее присутствии для изображений, представляющих собой реальные снимки городской застройки, в которые была внесена погрешность их привязки друг к другу.

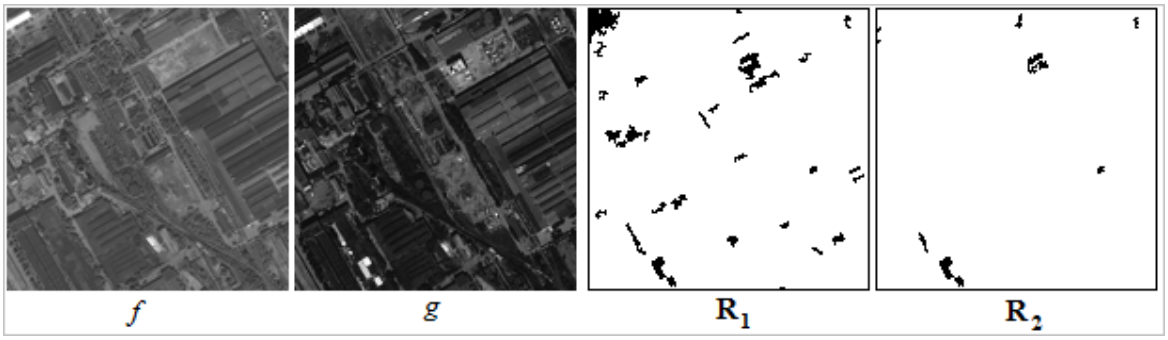

Рис. 3.10. Сравнение результатов работы алгоритма поиска структурных различий изображений f и g на основе морфологического проектора при отсутствии локальной геометрической коррекции ($\mathbf{R}_1$) и в ее присутствии ($\mathbf{R}_2$). Размер сканирующего окна $d = 31$, порог $T = 40$.

Параметры выбраны следующим образом: функция штрафа $h(\cdot) \equiv 1$, $\delta = 30$ и $\varepsilon = 30$. Локальный сдвиг можно оценить по сдвигу ярко-белого дома в левом верхнем углу снимков. Сдвиг изображений был сделан вручную на три пикселя вправо и три пикселя вниз. Соответственно, эти же параметры использовались при локальной геометрической коррекции с помощью вычисления “степени биективности”. Можно видеть, что без применения геометрической коррекции алгоритм нашел очень много лишних объектов — по сути, структурным различием стал любой значительный перепад яркости на снимках. В то же время тот же алгоритм, использующий вычисление “степени биективности” для геометрического сопоставления локальных окон, обнаружил только структурные различия. Более подробное исследование влияния рассинхронизации входных изображений на результат работы алгоритмов поиска структурных различий изображений будет приведено в следующей главе.

В целом, метод вычисления “степени биективности” показал себя полезным вспомогательным инструментом при решении задачи поиска структурных различий. Однако его широкое практическое применение все же представляется затруднительным из-за необходимости настройки дополнительных параметров. В реальных задачах для привязки снимков друг к другу целесообразнее использовать специализированные комплексы ГИС.

3.6 Асимптотическая оценка времени работы алгоритмов

Одним из важных этапов анализа алгоритмов является асимптотическая оценка времени работы, что позволяет, узнав требуемые затраты вычислительных ресурсов, правильно подобрать систему, на которой алгоритм будет работать, а также сравнить алгоритмы между собой. Подобный анализ будет проведен на основе подхода, предлагаемого в [16] и использующего модель однопроцессорной машины с произвольным доступом (параллельные вычисления будут рассмотрены в следующем пункте).

Очевидно, что время работы будет зависеть от размера входных изображений: обозначим его $size[f, g] = W \cdot H = S$. Время работы алгоритма — это число операций, которые он выполняет. Для выполняемых алгоритмами операций запишем их стоимость (каждая операция, помимо самой себя, содержит и стоимость присвоения результата некоторой переменной):

- c_1: перебор индекса в цикле **for** и проверка условия выхода, при этом ограничения вида $i \mod W - d/2$ учитываться не будут, поскольку непосредственно в алгоритме, разумеется, данные величины заданы без необходимости проведения дополнительных вычислений, а здесь будут приведены в таком виде исключительно для удобства чтения;

- c_2: обращение к памяти, для которого мы будем объединять непосредственно обращение и вычисление координаты точки на изображении типа $x + W \cdot y$, для простоты полагая, что обращение к элементу изображения, лежащему за его границами, возвращает априори невозможное значение яркости, например, -1;

- сложение c_3, вычитание c_4, умножение c_5, деление c_6;

- c_7: пороговая обработка, т. е. проверка того, превосходит ли заданная величина некоторое пороговое значение T;

- c_8: проверка соотношения двух величин в условии **if**;

- c_9: экспоненцирование;

- c_{10}: получение случайного числа с равномерным распределением из заданного интервала: $Rand(r_1, r_2)$;

- c_{11}: определение наименьшего среди двух чисел: $\min(\cdot, \cdot)$.

Далее запишем алгоритмы поиска структурных различий, основанные на различных функциях преобразования яркости, с помощью псевдокода, напоминающего распространенные языки программирования (C, Pascal и т.д.). Для каждой строчки запишем стоимость операции c_k и количество раз, которое алгоритм выполняет ее. В конце пункта будет приведена таблица, сравнивающая количество операций для различных алгоритмов. Для простоты все методы будут приведены без симметризации.

Первым рассмотрим алгоритм, основанный на морфологическом проекторе.

Морфологический проектор	*стоимость*	*число раз*
1 **for** $i \leftarrow 1$ **to** $W \cdot H$	c_1	$S+1$
2 $f_c \leftarrow f[i]$	c_2	S
3 **for** $x \leftarrow i \mod W - d/2$ **to** $i \mod W + d/2$	c_1	$S(d+1)$
4 **for** $y \leftarrow \lfloor i/W \rfloor - d/2$ **to** $\lfloor i/W \rfloor + d/2$	c_1	$Sd(d+1)$
5 **if** $f[x + W \cdot y] = f_c$	$c_2 + c_8$	Sd^2
6 $sum_i \leftarrow sum_i + g[x + W \cdot y]$	$c_2 + c_3$	$\sum_{i=1}^{S} t_i$
7 $sum_w \leftarrow sum_w + 1$	c_3	$\sum_{i=1}^{S} t_i$
8 $R[i] \leftarrow (sum_i/sum_w - g[i] > T)?255 : 0$	$2c_2 + c_4 + c_6 + c_7$	S

Здесь циклы **for** выполняются на один раз больше, чем операции, содержащиеся в них, из–за наличия последней проверки, выводящей из цикла. Также важно отметить, что некоторые операции выполняются строго фиксированное число раз, но количество заходов в цикл **if** будет зависеть от конкретного изображения: для выбранной точки i обозначим его через t_i. Следовательно, для всего изображения это количество будет равно $\sum_{i=1}^{S} t_i$. Фактически, величина t_i есть ничто иное, как площадь уровня яркости, соответствующего центральной точке локального окна: $t_i = |L_f(x_c)|$.

Сложив все строки, получим итоговое время работы, зависящее от параметров S и d:

$$T(S,d) = c_1(S+1) + c_2 S + c_1 S(d+1) + c_1 Sd(d+1) + (c_2 + c_8)Sd^2 +$$

$$+(c_2 + c_3)\sum_{i=1}^{S} t_i + c_3 \sum_{i=1}^{S} t_i + (2c_2 + c_4 + c_6 + c_7)S =$$

$$= S((c_1 + c_2 + c_8)d^2 + 2c_1 d + (2c_1 + 3c_2 + c_4 + c_6 + c_7)) + (c_2 + 2c_3)\sum_{i=1}^{S} t_i + c_1.$$

Далее рассмотрим два случая. В первом изображение f является одноуровневым, и, следовательно, алгоритм будет заходить в цикл **if** всегда (другими словами, площадь уровня яркости всегда совпадает с площадью сканирующего окна, равной d^2). Тогда $\sum_{i=1}^{S} t_i = \sum_{i=1}^{S} d^2 = S \cdot d^2$, и время работы — в наихудшем случае:

$$T(S,d) = S((c_1 + 2c_2 + 2c_3 + c_8)d^2 + 2c_1 d + (2c_1 + 3c_2 + c_4 + c_6 + c_7)) + c_1.$$

Второй случай — наилучший — заключается в том, что уровень яркости, соответствующий центральной точке окна, всегда состоит только из самой центральной точки (в частности, это

случай, когда входные изображения являются сильно зашумленными). Тогда $\sum\limits_{i=1}^{S} t_i = \sum\limits_{i=1}^{S} 1 = S$, и время работы:

$$T(S,d) = S((c_1 + c_2 + c_8)d^2 + 2c_1 d + (2c_1 + 4c_2 + 2c_3 + c_4 + c_6 + c_7)) + c_1.$$

В обоих крайних случая получился порядок роста $T(S,d) = \Theta(Sd^2)$. Интересно отметить, что алгоритм будет работать дольше в случае, когда изображения содержат мало шума, и при этом качество работы будет выше, а при сильном зашумлении время работы уменьшается, но вместе с ним и качество нахождения структурных различий.

Далее приведем такую же оценку для алгоритма, основанного на линейной функции преобразования яркости. Запишем его в виде псевдокода.

	Линейная функция	*стоимость*	*число раз*
1	**for** $i \leftarrow 1$ **to** $W \cdot H$	c_1	$S+1$
2	**for** $x \leftarrow i \mod W - d/2$ **to** $i \mod W + d/2$	c_1	$S(d+1)$
3	**for** $y \leftarrow \lfloor i/W \rfloor - d/2$ **to** $\lfloor i/W \rfloor + d/2$	c_1	$Sd(d+1)$
4	$f_{xy} \leftarrow f[x + W \cdot y]$	c_2	Sd^2
5	$g_{xy} \leftarrow g[x + W \cdot y]$	c_2	Sd^2
6	$sum_{x1} \leftarrow sum_{x1} + f_{xy}$	c_3	Sd^2
7	$sum_{y1} \leftarrow sum_{y1} + g_{xy}$	c_3	Sd^2
8	$sum_{xy} \leftarrow sum_{xy} + f_{xy} \cdot g_{xy}$	$c_3 + c_5$	Sd^2
9	$sum_{x2} \leftarrow sum_{x2} + f_{xy} \cdot f_{xy}$	$c_3 + c_5$	Sd^2
10	$sum_w \leftarrow sum_w + 1$	c_3	Sd^2
11	$k \leftarrow \frac{sum_w \cdot sum_{xy} - sum_{x1} \cdot sum_{y1}}{sum_w \cdot sum_{x2} - sum_{x1} \cdot sum_{x1}}$	$2c_4 + 4c_5 + c_6$	S
12	$b \leftarrow (sum_{y1} - k \cdot sum_{x1})/sum_w$	$c_4 + c_5 + c_6$	S
13	$R[i] \leftarrow (kf[i] + b - g[i] > T)?255 : 0$	$3c_2 + c_3 +$ $+c_4 + c_5 + c_7$	S

В данном алгоритме отсутствуют условные циклы, поэтому можно сразу же записать формулу для времени работы:

$$T(S,d) = c_1(S+1) + c_1 S(d+1) + c_1 Sd(d+1) + 2c_2 Sd^2 + 2c_3 Sd^2 + 2(c_3 + c_5)Sd^2 + c_3 Sd^2 +$$

$$+(2c_4 + 4c_5 + c_6)S + (c_4 + c_5 + c_6)S + (3c_2 + c_3 + c_4 + c_5 + c_7)S =$$

$$= S((c_1 + 2c_2 + 5c_3 + 2c_5)d^2 + 2c_1 d + (2c_1 + 3c_2 + c_3 + 4c_4 + 6c_5 + 2c_6 + c_7)) + c_1.$$

Таким образом, порядок роста $T(S,d) = \Theta(Sd^2)$.

Алгоритм, основанный на квадратичной функции преобразования яркости, в целом похож на предыдущий.

Квадратичная функция	*стоимость*	*число раз*
1 **for** $i \leftarrow 1$ **to** $W \cdot H$	c_1	$S+1$
2 **for** $x \leftarrow i \mod W - d/2$ **to** $i \mod W + d/2$	c_1	$S(d+1)$
3 **for** $y \leftarrow \lfloor i/W \rfloor - d/2$ **to** $\lfloor i/W \rfloor + d/2$	c_1	$Sd(d+1)$
4 $f_{xy} \leftarrow f[x + W \cdot y]$	c_2	Sd^2
5 $g_{xy} \leftarrow g[x + W \cdot y]$	c_2	Sd^2
6 $s_x \leftarrow sum_{x1} + f_{xy}$	c_3	Sd^2
7 $s_{x2} \leftarrow sum_{x2} + f_{xy}^2$	$c_3 + c_5$	Sd^2
8 $s_{x3} \leftarrow sum_{x3} + f_{xy}^3$	$c_3 + 2c_5$	Sd^2
9 $s_{x4} \leftarrow sum_{x4} + f_{xy}^4$	$c_3 + 3c_5$	Sd^2
10 $s_y \leftarrow sum_{y1} + g_{xy}$	c_3	Sd^2
11 $s_{xy} \leftarrow sum_{xy} + f_{xy} \cdot g_{xy}$	$c_3 + c_5$	Sd^2
12 $s_{x2y} \leftarrow sum_{x2y} + f_{xy}^2 \cdot g_{xy}$	$c_3 + 2c_5$	Sd^2
13 $s_w \leftarrow sum_w + 1$	c_3	Sd^2
14 $a \leftarrow \frac{(s_w s_{x2} - s_x)(s_w s_{x2y} - s_{x2} s_y) - (s_w s_{xy} - s_y)(s_w s_{x3} - s_x s_{x2})}{(s_w s_{x2} - s_x)(s_w s_{x4} - s_{x2}^2) - (s_w s_{x3} - s_{x2})(s_w s_{x3} - s_x s_{x2})}$	$10c_4 +$ $+12c_5 + c_6$	S
15 $b \leftarrow \frac{s_w s_{xy} - s_y - a(s_w s_{x3} - s_{x2})}{s_w s_{x2} - s_x}$	$4c_4 + 4c_5 + c_6$	S
16 $c \leftarrow (s_y - a s_{x2} - b s_x)/s_w$	$2c_4 + 2c_5 + c_6$	S
17 $f_c \leftarrow f[i]$	c_2	S
18 $R[i] \leftarrow (a f_c^2 + b f_c + c - g[i] > T)?255 : 0$	$2c_2 + 2c_3 +$ $+c_4 + 3c_5 + c_7$	S

Складывая стоимости по строкам, получаем:

$$T(S,d) = S((c_1 + 2c_2 + 8c_3 + 9c_5)d^2 + 2c_1 d +$$

$$+(2c_1 + 3c_2 + 2c_3 + 15c_4 + 21c_5 + 3c_6 + c_7)) + c_1.$$

Порядок роста такой же, как и для предыдущих алгоритмов: $T(S,d) = \Theta(Sd^2)$.

Далее рассмотрим алгоритм, основанный на регуляризованном варианте морфологического проектора.

Рег. морфологический проектор	*стоимость*	*число раз*
1 **for** $i \leftarrow 1$ **to** $W \cdot H$	c_1	$S+1$
2 **for** $j \leftarrow 1$ **to** I	c_1	$(I+1)S$
3 **for** $x \leftarrow i \mod W - d/2$ **to** $i \mod W + d/2$	c_1	$IS(d+1)$
4 **for** $y \leftarrow \lfloor i/W \rfloor - d/2$ **to** $\lfloor i/W \rfloor + d/2$	c_1	$ISd(d+1)$
5 $w_0 \leftarrow \exp(-(f[x+Wy]-j)^2/\sigma_c^2)\times$	$2c_2 + c_4 +$	ISd^2
$\times G_d(x - i \mod W, y - \lfloor i/W \rfloor)$	$+3c_5 + c_6 + c_9$	
6 $h \leftarrow h + w_0 \cdot g[x+Wy]$	$c_2 + c_3 + c_5$	ISd^2
7 $w \leftarrow w + w_0$	c_3	$256Sd^2$
8 $w_1 \leftarrow \exp(-\frac{(f[(i \mod W)+W(\lfloor i/W \rfloor)]-j)^2}{\sigma_c^2})\times$	$2c_2 + c_4 +$	IS
$\times G_d(\lfloor d/2 \rfloor, \lfloor d/2 \rfloor)$	$+3c_5 + c_6 + c_9$	
9 $sum_i \leftarrow sum_i + w_1 h/w$	$c_3 + c_5 + c_6$	IS
10 $sum_w \leftarrow sum_w + w_1$	c_3	IS
11 $R[i] \leftarrow (sum_i/sum_w - g[i] > T)?255 : 0$	$2c_2 + c_4 + c_6 + c_7$	S

Здесь I — количество допустимых уровней яркости исходных изображений (как правило, 256), $G(\cdot)$ понимается двумерное дискретное Гауссово распределение:

$$G_d(x,y) = \exp\Big(-\frac{(x-\lfloor d/2 \rfloor)^2 + (y-\lfloor d/2 \rfloor)^2}{\sigma_c^2}\Big),$$

σ_c — параметр сглаживания. Время работы:

$$T(S,d,I) = S(I(c_1 + 3c_2 + 2c_3 + c_4 + 4c_5 + c_6 + c_9)d^2 + 2Ic_1 d +$$

$$+((2I+2)c_1 + (2I+2)c_2 + 2Ic_3 + (I+1)c_4 + 4Ic_5 + (I+1)c_6 + c_7 + Ic_9)) + c_1.$$

Порядок роста теперь также зависит от количества допустимых уровней яркости I: $T(S,d,I) = \Theta(ISd^2)$.

Последний алгоритм — основанный на глобальной оптимизации. Как было указано в п. 3.4 данный метод был реализован с помощью алгоритма Метрополиса, включающего в себя параметр, отвечающий за количество итераций обработки изображения. Обозначим его N_e и будем считать равным 5000. Также следует напомнить, что в данном алгоритме параметр d играет роль размера окрестности точки, которая, как правило, меньше, чем размер сканирующего окна для предыдущих алгоритмов.

Глобальная оптимизация	*стоимость*	*число раз*
1 **for** $n \leftarrow 1$ **to** N_e	c_1	$N_e + 1$
2 **for** $i \leftarrow 1$ **to** $W \cdot H$	c_1	$(S+1)N_e$
3 $f_c \leftarrow f[i]$	c_2	SN_e
4 $g_c \leftarrow g[i]$	c_2	SN_e
5 $f'_c \leftarrow f'[i]$	c_2	SN_e
6 $r \leftarrow Rand(1, I)$	c_{10}	SN_e
7 **for** $x \leftarrow i \mod W - d/2$ **to** $i \mod W + d/2$	c_1	$S(d+1)N_e$
8 **for** $y \leftarrow \lfloor i/W \rfloor - d/2$ **to** $\lfloor i/W \rfloor + d/2$	c_1	$Sd(d+1)N_e$
9 $\delta \leftarrow \lvert f[x + Wy] - f_c \rvert$	$c_2 + c_4$	Sd^2N_e
10 **if** $\delta < \alpha$	c_8	Sd^2N_e
11 $g_{xy} \leftarrow g[x + W \cdot y]$	c_2	$\sum_{n=1}^{N_e}\sum_{i=1}^{S} t_{n,i}$
12 $u_1 \leftarrow u_1 + \min((g_{xy} - r)^2, \alpha^2)$	$c_3 + c_4 +$ $+2c_5 + c_{11}$	$\sum_{n=1}^{N_e}\sum_{i=1}^{S} t_{n,i}$
13 $u_2 \leftarrow u_2 + \min((g_{xy} - f'_c)^2, \alpha^2)$	$c_3 + c_4 +$ $+2c_5 + c_{11}$	$\sum_{n=1}^{N_e}\sum_{i=1}^{S} t_{n,i}$
14 $p \leftarrow \exp(-\frac{(r-f'_c)(r-2g_c+f'_c)+\lambda(u_1-u_2)}{Temp})$	$2c_3 + 4c_4 +$ $+2c_5 + c_6 + c_9$	SN_e
15 **if** $Rand(0,1) \le p$	$c_8 + c_{10}$	SN_e
16 $f'[i] = r$	c_2	$\sum_{n=1}^{N_e}\sum_{i=1}^{S} t'_{n,i}$
17 $Temp \leftarrow Temp \times 0.999$	c_5	N_e
18 **for** $i \leftarrow 1$ **to** $W \cdot H$	c_1	$S + 1$
19 $R[i] \leftarrow (\lvert f'[i] - g[i] \rvert > T)?255 : 0$	$2c_2 + c_4 + c_7$	S

В этом алгоритме два цикла **if**, поэтому вновь рассмотрим два случая: наилучший и наихудший. Для первого найдем наименьшее количество вхождений в циклы. Проверка $\delta < \alpha$ в строчке 10, очевидно, выполняется как минимум один раз — для центральной точки сканирующего окна, независимо от значения параметра α. Поэтому в “лучшем” случае (когда изображение содержит сильный шум из–за чего величина δ значительна)

$$\sum_{n=1}^{N_e}\sum_{i=1}^{S} t_{n,i} = \sum_{n=1}^{N_e}\sum_{i=1}^{S} 1 = SN_e,$$

а в “худшем” (слабоменяющееся или однотонное изображение)

$$\sum_{n=1}^{N_e}\sum_{i=1}^{S} t_{n,i} = \sum_{n=1}^{N_e}\sum_{i=1}^{S} d^2 = Sd^2N_e.$$

Выполнение второго условия зависит от случайного числа $Rand(0,1)$. Здесь “лучший” случай соответствует ситуации, когда оно всегда равно 1, а $p < 1$ и тогда алгоритм никогда не заходит внутрь цикла, а “худший” — ситуации, когда случайное число всегда равно 0 и

$$\sum_{n=1}^{N_e}\sum_{i=1}^{S} t'_{n,i} = \sum_{n=1}^{N_e}\sum_{i=1}^{S} 1 = SN_e.$$

Таким образом, в лучшем случае алгоритм работает за время

$$T(S,d,N_e) = S[N_e(c_1+c_2+c_4+c_8)d^2 + 2N_ec_1d+$$

$$+N_e(2c_1+4c_2+4c_3+6c_4+6c_5+c_6+c_8+c_9+2c_{10}+2c_{11})+c_1+2c_2+c_4+c_7] + 2(N_e+1)c_1 + N_ec_5,$$

а в худшем — за

$$T(S,d,N_e) = S[N_e(c_1+2c_2+2c_3+3c_4+4c_5+c_8+2c_{11})d^2 + 2N_ec_1d+$$

$$+N_e(2c_1+4c_2+2c_3+4c_4+2c_5+c_6+c_8+c_9+2c_{10})+c_1+2c_2+c_4+c_7] + 2(N_e+1)c_1 + N_ec_5.$$

Здесь порядок роста теперь также зависит от количества итераций алгоритма N_e: $T(S,d,N_e) = \Theta(N_eSd^2)$.

Далее объединим в таблицу количество операций каждого вида по всем алгоритмам. Результат представлен в табл. 3.1 (левая колонка), для алгоритмов, содержащих условный цикл **if** количество операций приведено в худшем случае. Алгоритмы удобно разбить на две группы: более быстрые — основанные на морфологическом проекторе, линейных и квадратичных функциях, и более вычислительно сложные — основанные на регуляризованном морфологическом проекторе и глобальной оптимизации. В каждой группе наблюдается последовательное увеличение числа операций, и, следовательно, времени работы алгоритмов.

Для большей наглядности вычислим количество операций для изображения размером 1000×1000 пикселей и оптимальных параметров размера сканирующего окна (окрестности пикселя), определению которых посвящена следующая глава (стандартное отклонение шума на изображениях полагается равным $\sigma = 14$):

Т а б л и ц а 3.1

Количество выполняемых алгоритмами поиска структурных различий операций: общая формула и для фиксированных параметров и изображения размером 1000×1000**,** $\times 10^6$

оп-я	*ст-ь*	Морф. пр.		Лин. ф-я		Кв. ф-я	
for	c_1	$S(d^2+2d+2)+1$	577	$S(d^2+2d+2)+1$	677	$S(d^2+2d+2)+1$	577
$[\cdot]$	c_2	$S(2d^2+3)$	1061	$S(2d^2+3)$	1253	$S(2d^2+3)$	1061
“+”	c_3	$2Sd^2$	1058	$S(5d^2+1)$	3126	$S(8d^2+2)$	4234
“-”	c_4	S	1	$4S$	4	$16S$	16
“×”	c_5	0	0	$S(2d^2+6)$	1256	$S(9d^2+21)$	4782
“/”	c_6	S	1	$2S$	2	$3S$	3
T	c_7	S	1	S	1	S	1
if	c_8	Sd^2	529	0	0	0	0
$\exp(\cdot)$	c_9	0	0	0	0	0	0
$Rand(\cdot,\cdot)$	c_{10}	0	0	0	0	0	0
$\min(\cdot,\cdot)$	c_{11}	0	0	0	0	0	0

оп-я	*ст-ь*	Рег. морф. пр.		Глоб. оптим	
for	c_1	$S(Id^2+2Id+2I+2)+1$	173 314	$S(N_ed^2+2N_ed+2N_e+1)+2N_e+2$	1 285 001,1
$[\cdot]$	c_2	$S(3Id^2+I+2)$	480 258	$S(2N_ed^2+4N_e+2)$	2 270 002
“+”	c_3	$S(2Id^2+2I)$	320 512	$S(2N_ed^2+2N_e)$	2 260 000
“-”	c_4	$S(Id^2+I+1)$	160 257	$S(3N_ed^2+4N_e+1)$	3 395 001
“×”	c_5	$S(4Id^2+4I)$	641 024	$S(4N_ed^2+2N_e)+N_e$	4 510 000,05
“/”	c_6	$S(Id^2+2I+1)$	160 513	SN_e	5000
T	c_7	S	1	S	1
if	c_8	0	0	$S(N_ed^2+N_e)$	1 130 000
$\exp(\cdot)$	c_9	$IS(d^2+1)$	160 256	SN_e	5000
$Rand(\cdot,\cdot)$	c_{10}	0	0	$2SN_e$	10 000
$\min(\cdot,\cdot)$	c_{11}	0	0	$2SN_ed^2$	2 250 000

- морфлогический проектор: $d = 23$ по формуле (4.3);

- линейная функция: $d = 25$ по формуле (4.4);

- квадратичная функция: $d = 23$ по формуле (4.5);

- регуляризованный вариант морфологического проектора: $d = 25$ по формуле (4.6), также $I = 256$;

- глобальная оптимизация: $d = 15$, также $N_e = 5000$.

Результаты приведены в таблице 3.1 (правая колонка). В итоге по каждой операции наблюдается рост числа вызовов, за исключением проверки условий в цикле **if**, отсутствующей в трех алгоритмах, и операции деления, максимальное количество которой приходится на алгоритм с регуляризованным вариантом морфологического проектора.

В пункте 3.8 будет приведено сравнение времени работы алгоритмов в секундах.

3.7 Программный комплекс для поиска структурных различий

Задача поиска структурных различий состоит в нахождении изменений типа появления / исчезновения объектов на паре разновременных снимков одного и того же участка земной поверхности. Однако на практике такой поиск является лишь первым шагом, за которым следует обработка и анализ полученных результатов. В пункте 1.2 приводились примеры задач анализа движения, в которых поиск различий служил предобработкой изображений, уменьшающей объем входных данных. Применительно к задачам картографии наибольший интерес представляют не структурные различия сами по себе, а их площадь — абсолютная в пикселях и вычисленная относительно номенклатурных листов [37], покрывающих исходные снимки. Подобная информация важна для своевременного обновления топографических карт: сначала по крупномасштабному снимку проводится оценка площади изменений, и если она значительна, заказываются более точные снимки изменившихся участков.

В данном пункте приводится описание программного комплекса, решающего задачу поиска структурных различий и состоящего из двух блоков:

1. Программный модуль, интегрированный в ГИС ENVI и разработанный с помощью встроенного в ENVI языка программирования IDL.

2. Модуль поиска структурных различий, представляющий собой динамически подключаемую библиотеку, написанную на MS Visual C++ 2008 Express Edition.

Программный модуль отвечает за загрузку и сохранение изображений, а также за визуализацию результата. Входными данными для алгоритма являются спутниковые снимки, как

правило, большого размера и требующие некоторой специальной предварительной обработки типа координатной привязки углов, что удобнее всего делать в разработанных для этих целей ГИС, среди которых была выбрана ГИС ENVI. В ENVI встроен язык программирования IDL, однако, он основан на вызове готовых процедур и не предназначен для реализации на нем сложных и нестандартных алгоритмов, поэтому основная вычислительная часть — поиск структурных различий и обработка результирующего изображения — была сделана в виде динамически подключаемой библиотеки. На рисунке 3.11 представлена общая схема разработанного программного комплекса.

Рис. 3.11. Общая схема программного комплекса, предназначенного для поиска структурных различий.

Далее подробно опишем порядок работы программного комплекса. Загрузка входных снимков осуществляется встроенными средствами ENVI, что позволяет решить две задачи: обеспечить широкий выбор форматов входных снимков (ENVI, будучи ГИС, работает со всеми основными типами файлов) и выполнить необходимую предобработку снимков (географическая привязка углов снимков и их выравнивание по размеру при необходимости). Полученная в результате пара геометрически выровненных и имеющих одинаковый размер снимков, а также файл с геопривязкой, которая понадобится в дальнейшем, передаются на вход модуля поиска структурных различий.

Модуль поиска структурных различий реализован в виде динамически подключаемой библиотеки (.dll), написанной на языке C++. Данный модуль содержит две части: собственно поиск различий и блок перевода контуров в текстовый формат. Поиск различий осуществляется с помощью алгоритма (пункт 3.1), основанного на регуляризованном морфологическом проекторе (п. 3.2.1). Поскольку требования к программному комплексу подразумевали его использование для крупномасштабных снимков, то размер сканирующего окна по результатам исследования

из пункта 4.4 (см. след. главу) был выбран $d = 21$, чтобы наилучшим образом соответствовать размерам присутствующих на снимках объектов (уже не отдельных строений, а целых кварталов, значительных участков леса и т. д.). Величина яркостного сглаживания была установлена равной $\sigma_c = 2$ (см. п. 4.4). Полученное разностное изображение возвращается ENVI для визуализации результата.

Далее следует этап пороговой обработки разностного изображения, которая осуществляется средствами ENVI, позволяющими также построить контуры вокруг найденных областей. Эти контуры накладываются на исходные снимки, ограничивая найденные области структурных различий (рис. 3.12). Выбор порогового значения осуществляется вручную с помощью полосы прокрутки (trackbar'a), чтобы пользователи имели возможность выбирать уровень изменений, существенных для них.

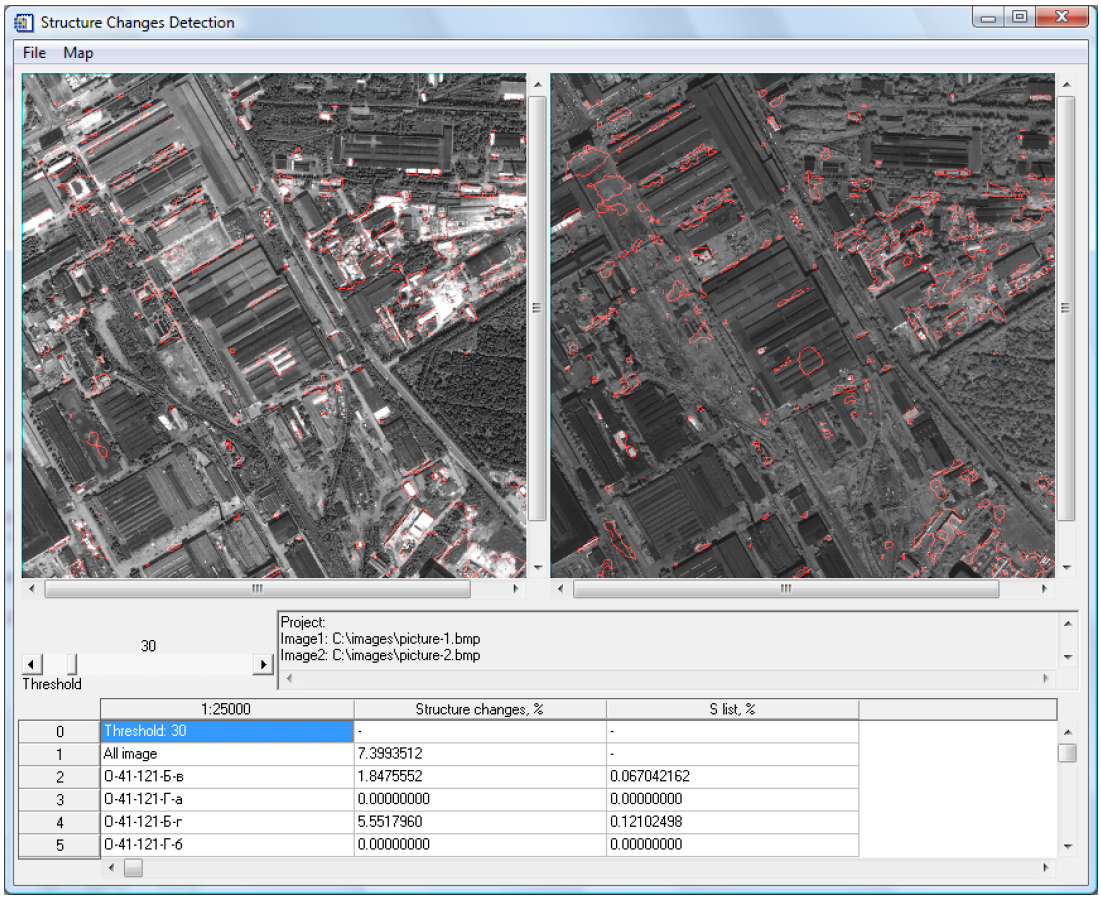

Рис. 3.12. Основное окно программного комплекса поиска структурных различий: сверху входные изображения с контурами найденных областей структурных различий, далее идет блок выбора порога и информация о входных снимках, внизу — таблица оценки площади изменений по номенклатурным листам, покрывающим снимки.

Для сохранения текущих настроек в программном комплексе используется механизм проектов пользователя. Проект содержит в себе конфигурационный файл, результирующее разностное изображение и, по необходимости, входные снимки. Это делается для того, чтобы не повторять процедуру предобработки снимков при каждом использовании программы. Конфигурационный файл проекта содержит следующую информацию: значение порога, путь к исходным снимкам (если при сохранении проекта не было указано сохранять исходные изображения) или имена файлов (если исходные изображения были сохранены), а также специальный признак конвертации, необходимый только для .tiff файлов. Открытие проекта позволяет не пересчитывать разностное изображение, а сразу загружать результат работы алгоритма и выводить необходимую информацию о найденных структурных различиях.

Получив изображения с найденными структурными различиями, можно приступить к анализу их площадей для различных масштабов и номенклатурных листов. Карты различных территорий являются многолистными, где каждый лист ограничен меридианами и параллелями, протяженность которых зависит от масштаба карты. Наличие таких карт потребовало создания определенной системы нумерации листов, называемой номенклатурой. В основу номенклатуры карт положена международная разграфка карты масштаба 1:1000000. Более подробно об индексах номенклатуры листов различных масштабов можно прочитать в [37]. Для работы с номенклатурными листами потребуется файл с геоприязкой углов снимков: листы имеют зафиксированные координаты, и зная координаты геопривязки углов изображений, можно определить номенклатурные листы, покрывающие исходные снимки.

Анализ степени структурных различий изображений проводится для определенного масштаба, заданного пользователем. Результатом анализа площади является таблица, в которой приводятся данные о степени структурного различия в процентах (рис. 3.12): в первой строке для всего снимка, в следующих строках — для номенклатурных листов указанного масштаба, имеющих общую часть с анализируемыми снимками. Степень различия для всего изображения вычисляется как выраженное в процентах отношение суммы площадей найденных областей со структурными различиями к площади всего изображения. Для отдельного номенклатурного листа вычисляется отношение суммы площадей пересечений таких областей с листом либо к площади всего изображения (вторая колонка таблицы), либо к площади всего листа (третья колонка). Полученная таблица позволяет судить о площади различия либо для снимка (в целом или по отдельным номенклатурным листам), либо для каждого номенклатурного листа. Если степень различия для некоторого листа оказалась значительной, принимается решение об обновлении карты, соответствующей данному листу.

Результаты работы программного комплекса можно сохранить в следующих видах:

— как изображения, представляющие собой входные снимки с наложенными на них контурами, ограничивающими найденные области структурных различий;

— как log-файл, содержащий таблицу площадей структурных различий для всего изображения и каждого номенклатурного листа по отдельности;

— как текстовые файлы в формате .mif / .mid, содержащие найденные контуры.

Остановимся на текстовом представлении контуров. Сохранять каждый контур поточечно, разумеется, невозможно — размер итогового файла будет огромен. Поэтому изображение с результатом пороговой обработки сжимается в 10 раз, после чего его размер восстанавливается. Контуры полученного таким образом изображения содержат намного меньше точек, и именно они и сохраняются в текстовом формате. Поскольку сохранение контуров необходимо для возможности их использования в прикладных приложениях (том же ENVI, например), для записи был выбран формат обмена данными MapInfo, позволяющий сопоставлять графическим данным различную информацию, и состоящий из двух файлов. Файл .mif (*MapInfo Interchange Format*) содержит графическую информацию (в нашем случае — координаты найденных контуров), файл .mid (*MapInfo Data File*) содержит текстовую информацию. Более подробно о формате MapInfo можно посмотреть в [46].

Разработанный программный комплекс был интегрирован в среду визуализации и обработки данных дистанционного зондирования Земли ENVI и внедрен для использования в работу ОАО “НИиП центр Природа”.

3.8 Использование графических процессоров

Применение предложенных в данной главе алгоритмов поиска структурных различий требует достаточно значительных вычислительных ресурсов. В целом это связано даже не со сложностью самих методов: например, алгоритм на основе морфологического проектора можно оптимизировать для чрезвычайно быстрой работы. Более важное значение имеет объем входных данных: обрабатываемые снимки содержат порядка 10^8 пикселей, причем это — нижняя оценка. Каждый такой пиксель требует обработки локального окна, центром которого он является, после чего следует этап пороговой и постобработки полученных результатов. Все это приводит к тому, что решение задачи поиска структурных различий требует значительного количества времени. Конечно, в данной задаче изначально не ставится требований выполнения в реальном времени, поскольку даже сами снимки делаются с интервалом в несколько лет (чтобы накопить достаточный объем изменений). Тем не менее, современный уровень развития технологий позволяет существенно сократить время расчетов, что сделает программное приложение более

удобным для практического использования.

В качестве такой технологии была выбрана CUDA — архитектура, разработанная компанией NVIDIA для вычислений на многоядерных графических процессорах. Возникший в связи с этим термин *Graphics Processing Unit* (GPU) используется для подчеркивания того факта, что графический ускоритель можно использовать для решения широкого круга задач, а не только для работы с графикой. Использование GPU позволяет в десятки раз ускорить выполнение программ, хотя и требует некоторых усилий по их распараллеливанию. Также использование CUDA позволяет проводить расчеты на суперкомпьютерах, что будет сделано при проведении численного эксперимента в следующей главе. Более подробно о технологии CUDA можно прочитать в [3, 41].

Распараллеливание алгоритмов поиска структурных различий следует достаточно очевидной схеме: поскольку алгоритмы работают со сканирующим окном, причем вычисления для каждого такого окна проходят независимо, то вполне разумно разделить программу так, чтобы каждое окно обрабатывалось на одном из процессоров GPU. Для алгоритма на основе глобальной оптимизации схема иная: корректировка текущей конфигурации осуществляется путем случайного изменения яркости каждого пикселя. Соответственно, работа для каждого процессора заключается в изменении яркости пикселя и проверки необходимости смены конфигурации в связи с этим. Поскольку при таком подходе необходимо работать сразу с двумя конфигурациями — текущей и новой, которая заменит текущую по окончании обработки изображения, следует применять двойной буфер (иногда называемый пинг-понг буфером) [113].

В таблице 3.2 приводятся данные о сравнении времени расчетов в секундах алгоритмов поиска структурных различий на CPU и GPU для изображений размером 1000×1000 пикселей. Параметры выбраны следующим образом: стандартное отклонение шума на изображениях $\sigma = 10$; для алгоритмов, использующих сканирующее окно, его размер $d = 21$; параметр яркостного сглаживания для регуляризованного морфологического проектора $\sigma_c = 2$; для алгоритма на основе глобальной оптимизации выбраны следующие параметры: $d = 15$, $\alpha = 5$ и $\lambda = 1816$. Для сравнения использовались процессор Intel Core i7-2600 3.40 GHz и видеокарта NVIDIA GeForce GTX 550 Ti.

Приведенное в таблице 3.2 время относится исключительно к обработке снимков и построению результирующего разностного изображения, т. е. той части алгоритмов, которая подлежит распараллеливанию, без постобработки снимков. Можно видеть, что наибольшее время работы с большим запасом показал алгоритм, основанный на глобальной оптимизации. Подобный результат был ожидаем, в том числе из-за идеологии создания данного метода: приведение изображений друг к другу не по зафиксированной формуле, а наиболее оптимальным образом.

Таблица 3.2

Сравнение времени работы в миллисекундах алгоритмов поиска структурных различий для CPU и GPU

Алгоритм	CPU	GPU
Морфологический проектор	1264	78.1
Линейная функция	2808	91.3
Квадратичная функция	4383	99.5
Регуляризованный морфологический проектор	46894	7965.8
Глобальная оптимизация	21993584	429152.1

При сравнении расчетов на CPU и GPU можно отметить значительное — в 6–51 раз (в зависимости от алгоритма) — уменьшение времени работы, что позволяет сделать вывод о необходимости использования данной технологии для практического поиска структурных различий изображений.

3.9 Выводы по главе 3

В данной главе были приведены алгоритмы, разработанные для решения задачи поиска структурных различий. Два из них — морфологический проектор для цветных изображений (п. 3.2.2) и “степень биективности” (п. 3.5) имеют вспомогательное значение и в дальнейшем обсуждении участвовать не будут. Остальные пять алгоритмов имеют важное прикладное значение: алгоритмы, основанные на морфологическом проекторе (п. 3.2), его регуляризованном варианте (п. 3.2.1), классе линейных (п. 3.3.1) и квадратичных функций (п. 3.3.2), а также на глобальной оптимизации (п. 3.4). Отсюда следуют два важных вопроса: какой из алгоритмов лучше и при каких значениях параметров. Для ответа на эти вопросы в следующей главе приводится описание и результаты численного эксперимента по сравнению качества работы этих алгоритмов.

Помимо этого, было рассмотрено практическое применение предложенных методик: асимптотические оценки времени их работы, приведено описание программного комплекса, разработанного для поиска структурных различий на разновременных космических снимках, который был интегрирован в ENVI — программный комплекс визуализации и обработки данных дистанционного зондирования Земли.

Глава 4

Исследование алгоритмов поиска структурных различий

Глава 3 содержит алгоритмы, разработанные для поиска структурных различий изображений. Данная глава состоит из двух частей: эмпирического исследования каждого алгоритма по отдельности и сравнения результатов их работы.

Работа каждого алгоритма напрямую зависит от выбранного набора его параметров. Ключевое место среди них занимает оптимальный порог, однако оценка требуется и для всех остальных параметров (например, для размера сканирующего окна). Среди всех возможных значений параметров требуется выбрать те, при которых алгоритм показывает наилучшее качество работы. Здесь сразу возникает вопрос: какие показатели использовать для наиболее адекватного сравнения алгоритмов. В случае выбора оптимального порога таким показателем будет *percentage correct classification*, т. е. минимум суммы ошибок первого и второго рода. Но для оценки всех остальных параметров, а также для сравнения алгоритмов между собой такой критерий представляется слишком жестким и сильно зависящим от входных данных. Более подходящим критерием является анализ ROC-кривых [65] (ROC — *Receiver Operating Characteristic*, операционная характеристика приёмника), появившийся в теории обработки сигналов в середине прошлого века. На основе этих двух критериев будет проведено исследование алгоритмов с целью выявления оптимальных параметров их работы в зависимости от уровня шума изображений. Разумеется, для подобного исследования потребуется специальная методика, учитывающая особенности математической модели из главы 2.

Второй целью исследования является сравнение качества работы алгоритмов между собой с целью определения того, который можно было бы рекомендовать для практического применения. Получив выводы о параметрах и качестве работы алгоритмов, можно применить получен-

ную информацию для решения прикладной задачи, с которой и началась данная диссертация: поиск и оценка площади структурных различий для разновременных космических снимков земной поверхности (описание программного комплекса дано в предыдущей главе).

Первым пунктом в исследовании алгоритмов будет утверждение, касающееся оптимального порога алгоритмов, основанных на морфологическом проекторе, линейных и квадратичных функциях преобразования яркости в случае отсутствия шума.

4.1 Оптимальный порог в случае отсутствия шума

Как было указано выше, последний шаг алгоритма поиска структурных различий связан с выбором порога T. Очевидно, что он является параметром, зависящим от изображений f и g, и правильная оценка его существенным образом влияет на результаты работы алгоритма. В данном пункте исследуется вопрос выбора порога в следующем идеальном случае: пусть изображение f — однотонное, т. е. $f(\cdot) \equiv c_0$, а изображение g состоит из двух однотонных областей U и V, соответствующих объекту и фону. При этом яркости $g(U) = b + i$, $g(V) = b$, площадь U составляет aS пикселей, где a — параметр из отрезка $[0, 0.5]$. Область U является структурным различием. Без ограничения общности можно считать $c_0 = b = 0$. Для такой модели верна следующая теорема.

Теорема 4.1. *Рассмотрим некоторый класс функций $\mathbb{F} = \{F\colon \mathbb{R} \to \mathbb{R}\}$, содержащий все константы, т. е. функции вида $F \equiv c$ для всех c. Будем использовать функции этого класса для попиксельных преобразований яркости исходных изображений.*

Пусть характеристическая функция структурной эквивалентности ω равна нулю для изображений с одинаковой структурой, и бесконечности в противном случае. Тогда при решении задачи поиска структурных различий для описанного выше идеального случая выбор порога из отрезка $(ai, (1-a)i]$ на этапе пороговой обработки будет гарантировать нахождение всех структурных различий без ложных срабатываний. В частности, для любых значений a оптимальный порог будет равен $i/2$.

Доказательство. Пороговая обработка как последний этап задачи поиска структурных различий применяется к результирующему разностному изображению $R = |f' - g|$, где f' — изображение, совмещающее структуру первого исходного изображения f и яркости второго исходного изображения g. По формуле (2.1):

$$f' = \arg\min(\|r - g\|^2 + \omega(r; f) \mid r \in \Phi).$$

Так как f — однотонное изображение, а $\mathbb{F}$ содержит все константы, то $\{F(f)\colon F \in \mathbb{F}\}$ является набором однотонных изображений $r(\cdot) \equiv c$ для всех c. Следовательно $\omega(r; f) = 0$ для всех f. Тогда выражение для построения проекции примет вид

$$f' = \arg\min(\|r - g\|^2 \mid r \in \Phi \text{ и } r(\cdot) \equiv \text{const}).$$

Его решением будет

$$f' = \frac{\sum\limits_{x \in X} g(x)}{S} = \frac{aS \cdot i}{S} = ai.$$

Далее производятся построение разностного изображения $R(x) = |f'(x) - g(x)|$ и его пороговая обработка. Для пикселей области структурного различия U требуется, чтобы $|ai - i| \geq T$, где T — некоторое пороговое значение. Для пикселей фона V, не являющихся структурным различием, требуется, чтобы $|ai - 0| < T$. Имеем

$$0 < ai < T \leq (1 - a)i.$$

Таким образом, множество допустимых порогов непусто и представляет собой полуинтервал с серединой $0.5 \cdot i$ и радиусом $(0.5 - a) \cdot i$, т. е. порог $0.5 \cdot i$ является оптимальным для всех значений a. Симметризация результата не требуется, так как изображение f — однотонное, и, следовательно, преобразованное изображение g' также будет однотонным, их разность $|g'(x) - f(x)| \equiv 0$. Теорема доказана. ■

Следствием этой теоремы является оценка влияния контраста входных изображений на результат работы алгоритма: чем выше контраст (в теореме это величина i), тем шире интервал оптимального порога, и, следовательно, тем легче его подобрать.

Следствие 4.1. *Для характеристических функций структурной эквивалентности ω_π, ω_1 и ω_2 в отсутствии шума оптимальный порог равен $0.5 \cdot i$.*

Данное следствие, верное для простой модели, разумеется, невозможно использовать для реальных снимков, но зато это строгий теоретический вывод, который сложно повторить в общем случае. Исключение — приведенные в пункте 2.6 теоремы для алгоритма, основанного на морфологическом проекторе, у которых, тем не менее, есть существенный недостаток: их трудно применять на практике. Поэтому далее будут приведено более простое эмпирическое исследование алгоритмов и оценка их оптимальных параметров, для чего предлагается следующая методика проведения численного эксперимента.

4.2 Математическая модель для проведения численного эксперимента

В практических задачах на изображениях, помимо полезного сигнала, обязательно присутствует шум, влияние которого необходимо учитывать при проведении эксперимента. Однако здесь проявляется основная трудность проведения исследований, свойственная любой задаче обработки изображений: влияние исходных данных на результаты. Суть такова: шум на реальных снимках плохо укладывается в априорные модели (независимость, аддитивность, нормальность распределения), и тем более невозможно изменить его неопределенные параметры. Второй момент, вызванный влиянием шума, — затруднительно построить надежную ручную разметку присутствующих на снимках структурных различий.

В второй главе диссертации была предложена математическая модель структурных различий, подходящая для теоретического исследования алгоритмов. Однако, несмотря на ее близость к реальным изображениям, спутниковые снимки все же намного сложнее. Тем не менее, для проведения численного эксперимента можно расширить эту модель путем усложнения "фона" исходных изображений и добавляемых "объектов", в качестве которых теперь будут рассматриваться не уровни яркости, а фрагменты спутниковых снимков, что позволит учесть следующие соображения:

1. Сравнение алгоритмов необходимо проводить на реальных, а не на синтезированных изображениях. Работа каждого алгоритма будет оцениваться путем сравнения (с помощью анализа ROC-кривых) полученных результатов с ручной разметкой исходных изображений. Однако получение ручной разметки реальных снимков представляет собой сложную и трудоемкую задачу. Причиной является сложность дешифрирования таких снимков, даже в ручном режиме. Иными словами, на снимках могут (и будут) присутствовать области, которые нельзя однозначно отнести к областям структурных различий или областям их отсутствия.

2. Стандартным подходом к определению устойчивости работы алгоритма является оценка его работы при различных уровнях шума, для чего необходимы изображения с "нулевым" уровнем возмущения. Параметры присутствующего на реальных изображениях шума можно оценить лишь приближенно, и использование подобной оценки в качестве "нулевого" уровня, к которому будет добавляться новый шум с известными параметрами, приведет к получению недостаточно надежного аргумента функции качества работы алгоритма.

3. Результат работы каждого алгоритма зависит от выбранного набора его параметров. Все параметры можно разделить на две группы: базовые (присущие всем алгоритмам) — оптимальный порог и размер сканирующего окна (размер окрестности пикселя для алгоритма

на основе глобальной оптимизации) и специфические — параметры сглаживания в алгоритме на основе регуляризованного морфологического проектора и коэффициенты энергетической функции для подхода, основанного на глобальной оптимизации. Как уже говорилось, ключевую роль играет оптимальный порог, однако размер сканирующего окна имеет не менее важное значение, поскольку его выбор определяет точность локализации объектов определенного размера. На реальных изображениях варьировать размер объектов представляется возможным только изменением масштаба всего изображения, что может быть невыгодно с точки зрения ограничения объема используемой информации: изображение может стать слишком маленьким, чтобы результаты его обработки представляли интерес. Также важным моментом является типовой состав объектов, так как в конкретных практических приложениях важна работа алгоритма для различных типов сцен (городская застройка, пригород и т. д.). Таким образом, модель должна включать в себя возможность изменения типа и размеров объектов.

Построения пары изображений, которые станут входными данными для алгоритма поиска структурных различий, и изображения с ручной разметкой осуществляется следующим образом. Выбирается космический снимок некоторого участка земной поверхности. В зависимости от практических требований это может быть снимок городской застройки, пригород и т. д. Масштаб снимка варьируется, чтобы получить требуемый размер объектов. Далее этот снимок разделяется на две части: первый фрагмент и его копия становятся исходными изображениями, второй фрагмент будет донором для генерации структурных различий.

Поскольку исходные изображения изначально идентичны, то одинаковы и присутствующие на них уровни яркости, вызванные влиянием шума. Таким образом, эти уровни яркости становятся в некотором смысле “содержательными”, т. е. определяют “фон” на обоих изображениях. Применение алгоритмов поиска структурных различий к таким изображениям, очевидно, дает нулевое результирующее разностное изображение R.

Структурными различиями становятся объекты, вырезанные из второго фрагмента снимка и добавленные к исходным изображениям. При этом добавление объекта ко второму изображению эквивалентно его появлению; добавление к первому изображению — исчезновению объекта; добавление двух разных объектов к одной и той же области обоих снимков — изменению формы. Типы переносимых объектов также определяются требованиями задачи (например, дома, если это анализ городской застройки, участки лесного массива, если это задача мониторинга вырубки лесов и т. д.).

Таким образом строятся два изображения, содержащие необходимые структурные различия при “нулевом” уровне шума. Из них путем попиксельного вычитания получается изображение “ручной разметки” структурных различий, принимаемое за эталон. На рис. 4.1 представлены

Рис. 4.1. Фрагменты изображений городской застройки с "нулевым" уровнем шума для сравнения алгоритмов поиска структурных различий и их ручная разметка.

фрагменты полученных таким способом изображений. Далее к этим изображениям добавляется шум с известным распределением. В итоге получается вероятностная модель структурных различий, аналогичная модели из главы 2, но позволяющая реализовывать более сложные случаи, встречающиеся на практике. Использование данной модели позволяет создать пары изображений, в которых негативное влияние посторонних факторов будет сведено к минимуму, благодаря чему можно провести исследование и сравнение алгоритмов. В качестве методики оценки качества работы алгоритмов был выбран анализ ROC-кривых. В следующем пункте приводится его краткое описание.

4.3 Анализ ROC-кривых

Понятие ROC-кривых [65] (*Receiver Operating Characteristic*, операционная характеристика приёмника, иногда называют кривой ошибок) взято из методологии анализа качества приёма сигнала (*Signal Detection Analysis*), разработанной в середине прошлого века. Теория, стоящая за этим анализом (*Theory of Signal Detectability*, теория определимости сигнала), хотя и происходит первоначально из электроники и электротехники, может также быть успешно применена в области обработки изображений. ROC-кривая представляет собой график, отображающий соотношение между долей верных положительных классификаций от общего числа положительных классификаций (*true positive rate*, TPR, называемой чувствительностью алгоритма классификации) с долей ошибочных положительных классификаций от общего числа отрицательных классификаций (*false positive rate*, FPR; величина $1 - \mathrm{FPR}$ называется специфичностью алгоритма классификации) при варьировании порога решающего правила T. Формулы для расчета

верных и ошибочных положительных классификаций следующие:

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}, \qquad \mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}}, \tag{4.1}$$

где

— TP (*True Positives*) — верно классифицированные положительные примеры (так называемые истинно положительные случаи);

— TN (*True Negatives*) — верно классифицированные отрицательные примеры (истинно отрицательные случаи);

— FN (*False Negatives*) — положительные примеры, классифицированные как отрицательные (ошибка I рода). Это, так называемый, “пропуск”, когда интересующее нас событие ошибочно не обнаруживается (ложно отрицательные примеры);

— FP (*False Positives*) — отрицательные примеры, классифицированные как положительные (ошибка II рода). Это — ложное срабатывание, так как при отсутствии события ошибочно выносится решение о его присутствии (ложно положительные случаи).

Количественную интерпретацию ROC-кривой дает показатель AUC (*Area Under the Curve*, площадь под ROC-кривой) — площадь, ограниченная ROC-кривой и осью доли ложных положительных классификаций FPR. Чем выше показатель AUC, тем качественнее классификатор, при этом значение 0.5 соответствует случайному решению.

4.4 Оценка оптимальных параметров алгоритмов для зашумленных изображений

Определив критерий качества работы алгоритмов, можно провести численный эксперимент, который строится следующим образом. Для каждого значения параметра $\sigma \in [0, 25]$ с шагом 0.5 — стандартного отклонения добавленного шума — генерируется пара изображений по описанной в пункте 4.2 модели. К этим изображениям применяется алгоритм поиска структурных различий с одной из функций преобразования яркости (и определенным набором параметров самой функции). По полученному результирующему разностному изображению R и ручной разметке, перебирая все возможные значения порога, строится ROC-кривая. Критерием качества работы алгоритма считается площадь под ROC-кривой AUC.

Схемы серий испытаний для различных алгоритмов несколько отличаются, но для методов с морфологическим проектором, линейными и квадратичными функциями испытания проводятся по общей схеме:

1. Генерация изображений с добавлением шума, стандартное отклонение которого выбирается из интервала $\sigma \in [0, 25]$ с шагом 0.5.

2. Выбор размера сканирующего окна $d \in [11, 111]$ с шагом 2.

3. К полученной паре изображений применяется алгоритм с одной из функций преобразования яркости и выбранным параметром d. Результат работы алгоритма — результирующее разностное изображение R — после применения пороговой обработки сравнивается с ручной разметкой, на основе чего находятся значения TP, TN, FN, FP и вычисляется величина AUC. Также определяется оптимальное значение порога T_{opt}, дающее наименьшее значение суммы ошибок первого и второго рода.

4. Шаги 1–3 повторяются 10 раз с одинаковыми значениями параметров, и полученные результаты усредняются.

5. Перебирая все значения параметров, формируются таблицы результатов для выбранного варианта алгоритма $\mathrm{AUC}(\sigma, d)$ и $T_{opt}(\sigma, d)$, по которым строятся соответствующие графики, и для оптимального порога находится аппроксимирующая функция. Аппроксимация проводится с помощью метода наименьших квадратов полиномом пятой степени по переменной d и третей степени по σ:

$$T_{opt} = z_0 + \sum_{k=1}^{5} a_k d^k + \sum_{k=1}^{3} b_k \sigma^k + \sum_{k=1}^{5} c_k d^k \sigma + \sum_{k=1}^{5} e_k d^k \sigma^2 + \sum_{k=1}^{5} h_k d^k \sigma^3. \tag{4.2}$$

Для аппроксимации используется пакет OriginPro 8.0 [100].

6. Фиксируя значение σ, определяются значения d, соответствующие максимальному значению AUC. Полученные данные используются для построения эмпирической зависимости размера сканирующего окна d от величины σ. Аппроксимация проводится линейными функциями с помощью пакета OriginPro 8.0.

Ввиду своей крайне высокой вычислительной сложности, численный эксперимент проводился с использованием технологии CUDA (подробнее в пункте 3.8) и суперкомпьютера “Уран” ИММ УрО РАН [45]. Далее приводятся результаты экспериментов для алгоритмов с различными функциями преобразования яркости.

Морфологический проектор.

График функции $\mathrm{AUC}(\sigma, d)$ изображен на рис. 4.2 (1). Полученные коэффициенты эмпирической формулы (4.2) для оптимального порога представлены в таблице 4.1. Точность аппроксимации — величина, называемая коэффициентом детерминации $K_D(T_{opt}; d, \sigma)$ [1] и представляющая собой долю дисперсии зависимой переменной (T_{opt}), объясняемую рассматриваемой моделью (функцией 4.2): $K_D = 0$ означает полное отсутствие какой-либо связи между T_{opt} и переменными d, σ; $K_D = 1$ соответствует случаю строгой функциональной зависимости.

Таблица 4.1

Коэффициенты аппроксимации оптимального порога алгоритмов для формулы (4.2)

Коэффициент	Морфологический проектор	Линейная функция	Квадратичная функция
z_0	-0.77443	-0.97165	-1.80544
a_1	-0.04983	0.38224	0.53791
a_2	0.01612	-0.00495	-0.01365
a_3	$-4.05 \cdot 10^{-4}$	$-1.56 \cdot 10^{-5}$	$1.89 \cdot 10^{-4}$
a_4	$3.77 \cdot 10^{-6}$	$6.35 \cdot 10^{-7}$	$-1.44 \cdot 10^{-6}$
a_5	$-1.22 \cdot 10^{-8}$	$-2.87 \cdot 10^{-9}$	$4.62 \cdot 10^{-9}$
b_1	1.4981	1.88016	2.19544
b_2	-0.14954	-0.08289	-0.0995
b_3	0.00256	0.00151	$1.99 \cdot 10^{-3}$
c_1	0.06069	-0.07344	-0.12233
c_2	-0.051	$6.50 \cdot 10^{-4}$	$3.30 \cdot 10^{-3}$
c_3	$1.13 \cdot 10^{-4}$	$1.32 \cdot 10^{-5}$	$-4.83 \cdot 10^{-5}$
c_4	$-1.01 \cdot 10^{-6}$	$-2.35 \cdot 10^{-7}$	$3.87 \cdot 10^{-7}$
c_5	$3.21 \cdot 10^{-9}$	$9.63 \cdot 10^{-10}$	$-1.28 \cdot 10^{-9}$
e_1	0.00836	0.00981	0.01276
e_2	$-8.26 \cdot 10^{-5}$	$-2.07 \cdot 10^{-4}$	$-3.77 \cdot 10^{-4}$
e_3	$-6.86 \cdot 10^{-7}$	$1.62 \cdot 10^{-6}$	$5.74 \cdot 10^{-6}$
e_4	$1.45 \cdot 10^{-8}$	$-2.42 \cdot 10^{-9}$	$-4.51 \cdot 10^{-8}$
e_5	$-5.91 \cdot 10^{-11}$	$-1.46 \cdot 10^{-11}$	$1.42 \cdot 10^{-10}$
h_1	$-1.69 \cdot 10^{-4}$	$-2.33 \cdot 10^{-4}$	$-2.94 \cdot 10^{-4}$
h_2	$2.46 \cdot 10^{-6}$	$5.80 \cdot 10^{-6}$	$8.91 \cdot 10^{-6}$
h_3	$-7.46 \cdot 10^{-9}$	$-6.05 \cdot 10^{-8}$	$-1.34 \cdot 10^{-7}$
h_4	$-8.24 \cdot 10^{-11}$	$2.54 \cdot 10^{-10}$	$1.03 \cdot 10^{-9}$
h_5	$4.72 \cdot 10^{-13}$	$-2.61 \cdot 10^{-13}$	$-3.18 \cdot 10^{-12}$

В данном случае $K_D(T_{opt}; d, \sigma) = 0.99867$, на основе чего можно заключить, что представленная модель достаточно точно отражает реальную зависимость показателей. Полученная эмпирическая зависимость оптимального размера сканирующего окна d от величины σ изображена на рис. 4.3 (1) и имеет вид:

$$d = 0.47674 \cdot \sigma + 16.727. \tag{4.3}$$

Оценка точности: $K_D(d; \sigma) = 0.86845$. ROC-кривые представлены на рис. 4.4 (1).

Линейная функция преобразования яркости.

График функции $\mathrm{AUC}(\sigma, d)$ изображен на рис. 4.2 (2). Полученные коэффициенты эмпирической формулы (4.2) для оптимального порога представлены в соответствующей колонке таблицы 4.1. Точность аппроксимации: $K_D(T_{opt}; d, \sigma) = 0.99861$. Полученная эмпирическая зависимость оптимального размера сканирующего окна d от величины σ изображена на рис. 4.3 (2) и имеет вид:

$$d = 1.22824 \cdot \sigma + 8.07843. \tag{4.4}$$

Оценка точности: $K_D(d; \sigma) = 0.95703$. ROC-кривые представлены на рис. 4.4 (2).

Квадратичная функция преобразования яркости.

График функции $\mathrm{AUC}(\sigma, d)$ изображен на рис. 4.2 (3). Полученные коэффициенты эмпирической формулы (4.2) для оптимального порога представлены в соответствующей колонке таблицы 4.1. Точность аппроксимации: $K_D(T_{opt}; d, \sigma) = 0.99861$. Полученная эмпирическая зависимость оптимального размера сканирующего окна d от величины σ изображена на рис. 4.3 (3) и имеет вид:

$$d = 1.05448 \cdot \sigma + 8.40724. \tag{4.5}$$

Оценка точности: $K_D(d; \sigma) = 0.94504$. ROC-кривые представлены на рис. 4.4 (3).

Выводы.

В целом, результаты исследования для этих трех алгоритмов оказались близки. Глядя на графики $\mathrm{AUC}(\sigma, d)$ можно заметить, что наибольшие значения соответствуют $\sigma = 0$ (что вполне ожидаемо) и уменьшаются с ростом величины стандартного отклонения шума примерно по параболе. Однако, это уменьшение неравномерно: оно достаточно быстрое при малых и больших значениях d, но при этом более плавное между ними. Соответственно, графики имеют перегиб, который показывает наилучшие значения d для различных уровней шума. Аналогичный вывод можно сделать из рис. 4.4: при $d = 11$ ROC-кривые имеет площадь заметно меньше, чем при $d = 21$ или 31, а затем с ростом d ROC-кривые начинают удаляться от точки с координатами (0,1). Обнаруженные таким образом оптимальные значения параметра d использовались при построении зависимости $d(\sigma)$.

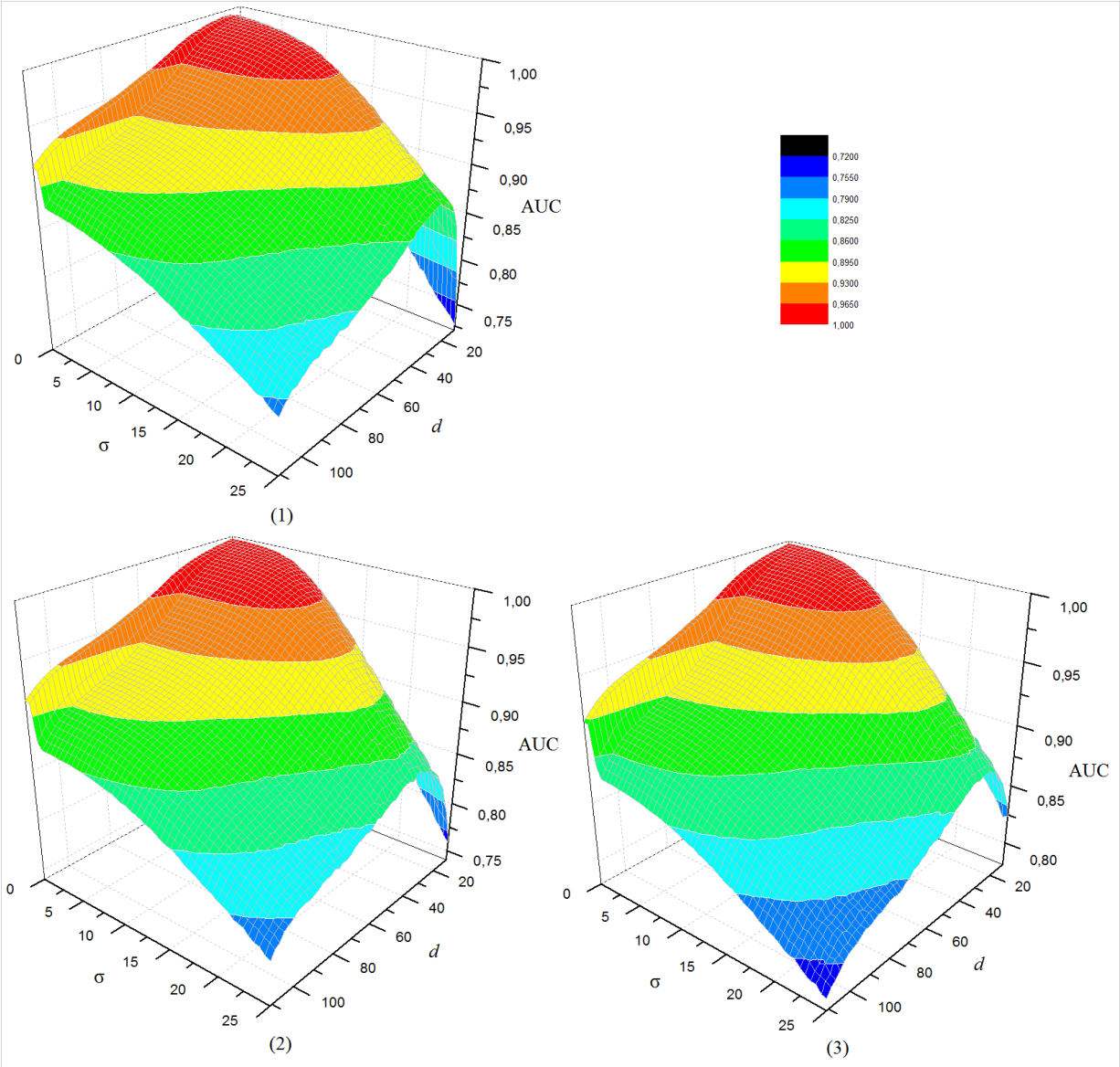

Рис. 4.2. Эмпирически полученные графики зависимости площади под ROC-кривой AUC от размера окна d и стандартного отклонения шума σ для алгоритма, основанного на:
(1) — морфологическом проекторе, (2) — линейной функции преобразования яркости,
(3) — квадратичной функции преобразования яркости.

Значение $d \approx 21$ оказалось наилучшим, что связано с размером объектов на снимке: для наилучшего обнаружения и локализации структурных различий линейные размеры объектов должны быть примерно вдвое меньше, чем размер сканирующего окна. При таком выборе объект, с одной стороны, не теряется на общем фоне (все такие алгоритмы работают на усреднении, и вклад малых областей может оказаться слишком незначительным), с другой стороны, объект полностью попадает “в кадр” (простой пример: внутренняя область даже самого явного структурного различия не будет найдена, если данный объект целиком покрывает сканирующее окно, — алгоритму становится просто не с чем его сравнивать). Но это рассуждение полностью верно лишь для идеальных незашумленных изображений. При увеличении уровня шума значительный размер сканирующего окна начинает играть роль “сглаживающего фильтра”, подавляя шум и тем самым способствуя повышению качества работы. Два этих ограничения задают некоторый интервал, в котором и лежит оптимальное значение параметра d. В следующем пункте будут приведены результаты применения полученной аппроксимации оптимального порога для модельных и для реальных изображений.

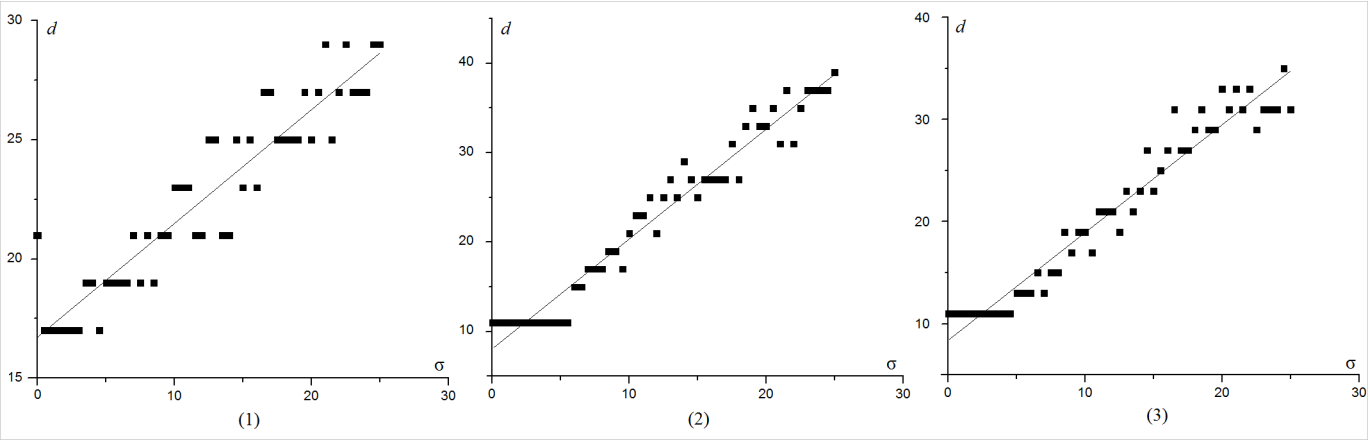

Рис. 4.3. Эмпирически полученные графики зависимости размера сканирующего окна d от величины стандартного отклонения шума σ для алгоритма, основанного на:
(1) — морфологическом проекторе, (2) — линейной функции преобразования яркости,
(3) — квадратичной функции преобразования яркости. Точками показаны результаты расчетов, линией — аппроксимирующая линейная функция.

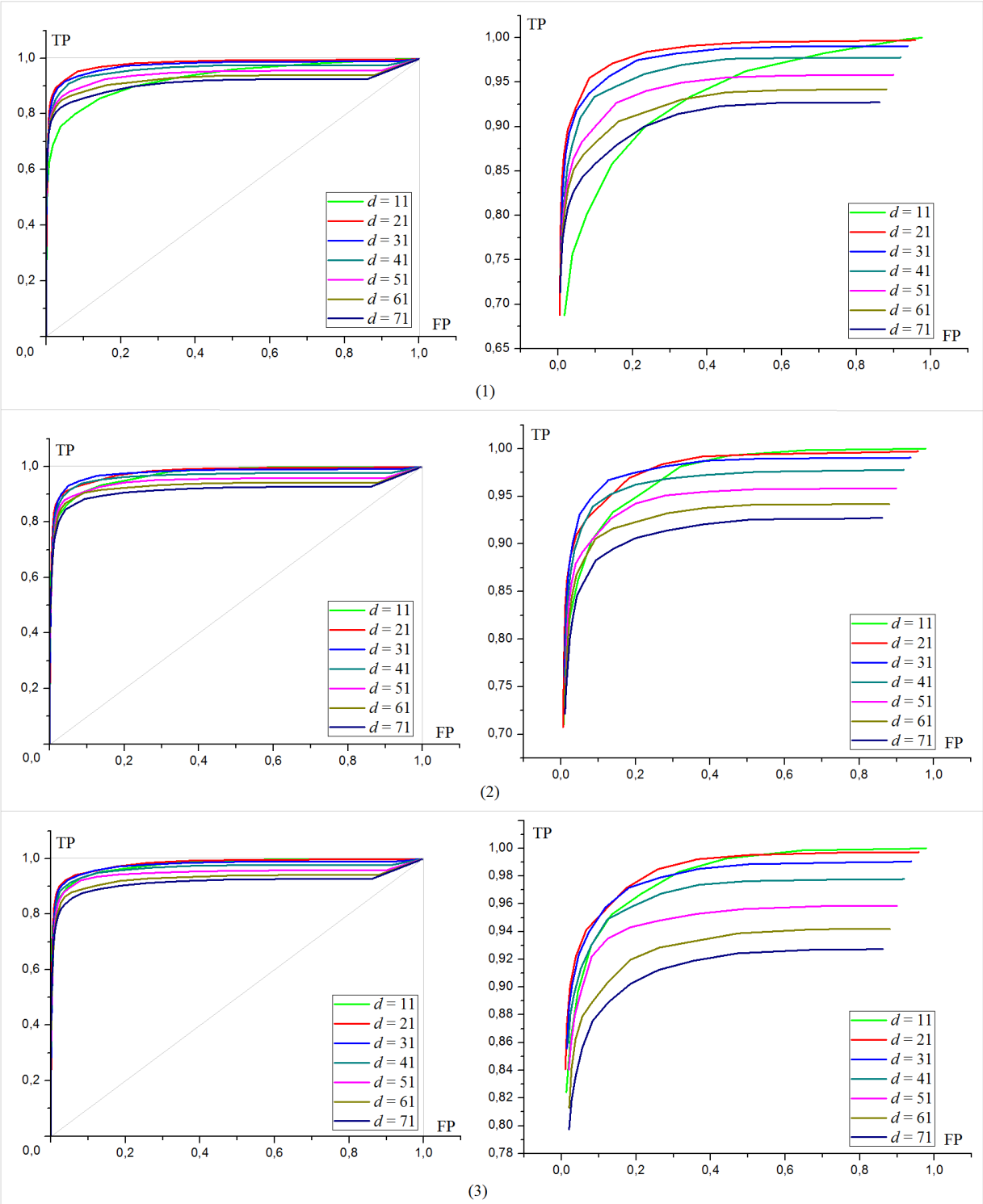

Рис. 4.4. Графики ROC-кривых для алгоритма, основанного на: (1) — морфологическом проекторе, (2) — линейной функции преобразования яркости, (3) — квадратичной функции преобразования яркости. Слева — полный график, справа — увеличенный фрагмент в месте перегиба функции.

Регуляризованный морфологический проектор.

Для алгоритма поиска структурных различий на основе регуляризованного морфологического проектора численный эксперимент проводится следующим образом. Как и раньше, для каждого значения параметра $\sigma \in [0, 25]$ с шагом 0.5 — стандартного отклонения добавленного шума — генерируется пара изображений по описанной в пункте 4.2 модели. К этим изображениям применяется алгоритм поиска структурных различий с определенным параметром яркостного сглаживания σ_c и размером сканирующего окна d. По полученному результирующему разностному изображению R и ручной разметке, перебирая все возможные значения порога, строится ROC-кривая. Критерием качества работы алгоритма считается площадь под ROC-кривой AUC. Схема испытаний следующая:

1. Генерация изображений с добавлением шума, стандартное отклонение которого выбирается из интервала $\sigma \in [0, 25]$ с шагом 0.5.

2. Выбор размера сканирующего окна $d \in [11, 71]$ с шагом 2 и параметра яркостного сглаживания $\sigma_c \in [1, 8]$ с шагом 1.

3. К полученной паре изображений применяется алгоритм с регуляризованным морфологическим проектором и выбранными параметрами d и σ_c. Результат работы алгоритма — результирующее разностное изображение R — после применения пороговой обработки сравнивается с ручной разметкой для нахождения значений TP, TN, FN, FP, и вычисляется величина AUC. Также определяется оптимальное значение порога T_{opt}, дающее наименьшее значение суммы ошибок первого и второго рода.

4. Шаги 1–3 повторяются 10 раз с одинаковыми значениями параметров, и полученные результаты усредняются.

5. Перебирая все значения параметров, формируются таблицы результатов $\mathrm{AUC}(\sigma, d, \sigma_c)$ и $T_{opt}(\sigma, d, \sigma_c)$. При фиксированном d выбираются значения σ_c, при которых величина AUC максимальна, и строится график $\sigma_c(d)$. Соответственно, можно упростить таблицы для AUC и T_{opt} так, чтобы они зависели только от двух параметров, тем самым приведя их к тому же виду, что и для предыдущих алгоритмов: $\mathrm{AUC}(\sigma, d)$ и $T_{opt}(\sigma, d)$. Уже по этим сокращенным таблицам также строятся соответствующие графики, и для оптимального порога находится аппроксимирующая функция. Аппроксимация проводится с помощью метода наименьших квадратов полиномом пятой степени по переменной d и третей степени по σ (4.2). Для аппроксимации используется пакет OriginPro 8.0.

6. Фиксируя значение σ, определяются значения d, соответствующие максимальному значению AUC. Полученные данные используются для построения эмпирической зависимости размера сканирующего окна d от величины σ. Аппроксимация проводится линейными функциями

Таблица 4.2

Коэффициенты аппроксимации оптимального порога алгоритма на основе регуляризованного морфологического проектора для формулы (4.2)

Коэффициент	z_0	a_1	a_2	a_3	a_4	a_5
Значение	9.3731	-1.27	0.07	$-1.53 \cdot 10^{-3}$	$1.65 \cdot 10^{-5}$	$-6.54 \cdot 10^{-8}$
Коэффициент	b_1	b_2	b_3	c_1	c_2	c_3
Значение	-1.21753	0.1268	-0.00381	0.465	$-2.69 \cdot 10^{-2}$	$6.59 \cdot 10^{-4}$
Коэффициент	c_4	c_5	e_1	e_2	e_3	e_4
Значение	$-7.43 \cdot 10^{-6}$	$3.17 \cdot 10^{-8}$	-0.03178	$1.97 \cdot 10^{-3}$	$-5.04 \cdot 10^{-5}$	$5.88 \cdot 10^{-7}$
Коэффициент	e_5	h_1	h_2	h_3	h_4	h_5
Значение	$-2.59 \cdot 10^{-9}$	$8.29 \cdot 10^{-4}$	$-5.16 \cdot 10^{-5}$	$1.35 \cdot 10^{-6}$	$-1.61 \cdot 10^{-8}$	$7.26 \cdot 10^{-11}$

с помощью пакета OriginPro 8.0.

Были получены следующие результаты:

— Оптимальное значение параметра яркостного сглаживания σ_c (т. е. такое σ_c, при котором для фиксированного d величина AUC будет максимальна) убывает с ростом d (рис. 4.5). Это связано с тем, что большая площадь сканирующего окна сама по себе играет роль сглаживания, и дополнительное размытие с помощью алгоритма не влияет на результат. В то же время при небольшом d (порядка 11) оптимальное значение $\sigma_c = 8$.

— График зависимости $\mathrm{AUC}(\sigma, d)$ (рис. 4.6) получился похожим на соответствующие графики для предыдущих алгоритмов, следовательно, можно провести те же рассуждения: аппроксимировать зависимость величины оптимального порога T_{opt} от σ и d по формуле (4.2). Полученные коэффициенты аппроксимирующей функции — в таблице 4.2; точность аппроксимации $K_D(T_{opt}; d, \sigma) = 0.99507$. Далее определяется оптимальное значение размера сканирующего окна d в зависимости от величины стандартного отклонения шума σ (рис. 4.7). Полученная формула линейной аппроксимации:

$$d = 1.13557 \cdot \sigma + 10.09955. \tag{4.6}$$

Оценка точности: $K_D(d; \sigma) = 0.94375$.

— Для получения зависимости качества работы алгоритма от величины σ_c на рис. 4.8 приведен график $\mathrm{AUC}(\sigma_c)$ при $d = 21$ и уровне шума $\sigma = 10$, достигающий максимума при $\sigma_c = 2$. ROC-кривые для различных значений d (при оптимальном выборе σ_c) представлены на рис. 4.9 (1). Можно наблюдать ту же картину, что и для предыдущих алгоритмов: наиболее близок к точке с координатами (0,1) график, соответствующий алгоритму с параметром $d = 21$.

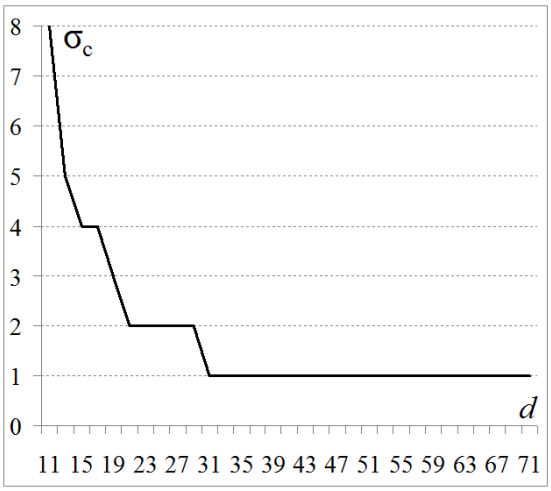

Рис. 4.5. График зависимости оптимальной (с точки зрения максимума AUC) величины σ_c от размера сканирующего окна d для алгоритма, основанного на регуляризованном морфологическом проекторе.

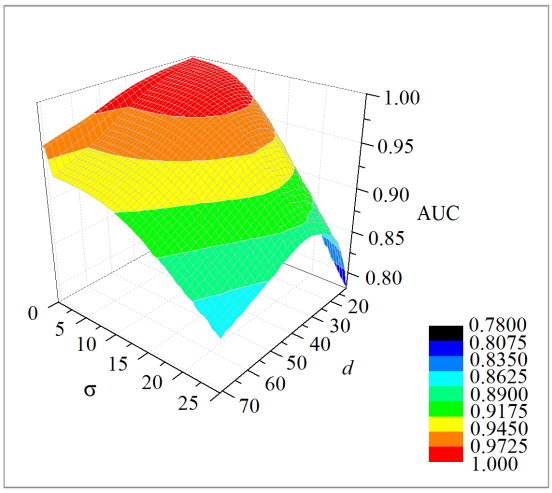

Рис. 4.6. Эмпирически полученный график зависимости площади под ROC-кривой AUC от размера окна d и стандартного отклонения шума σ для алгоритма, основанного на регуляризованном морфологическом проекторе.

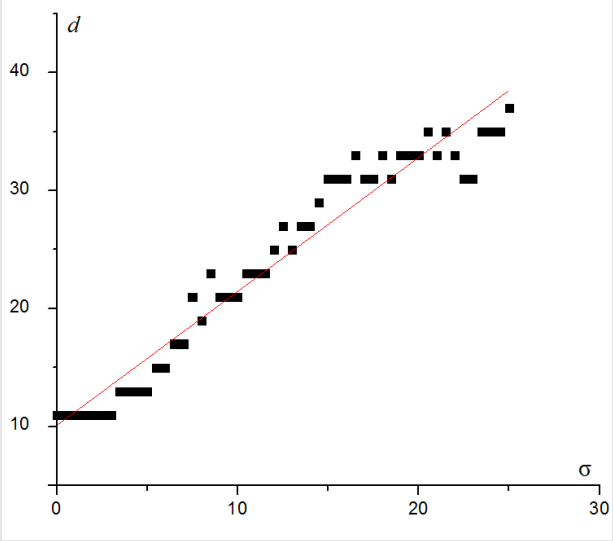

Рис. 4.7. Эмпирически полученный график зависимости размера сканирующего окна d от величины стандартного отклонения шума σ для алгоритма, основанного на регуляризованном морфологическом проекторе. Черными точками показаны результаты расчетов, красной линией — аппроксимирующая линейная функция.

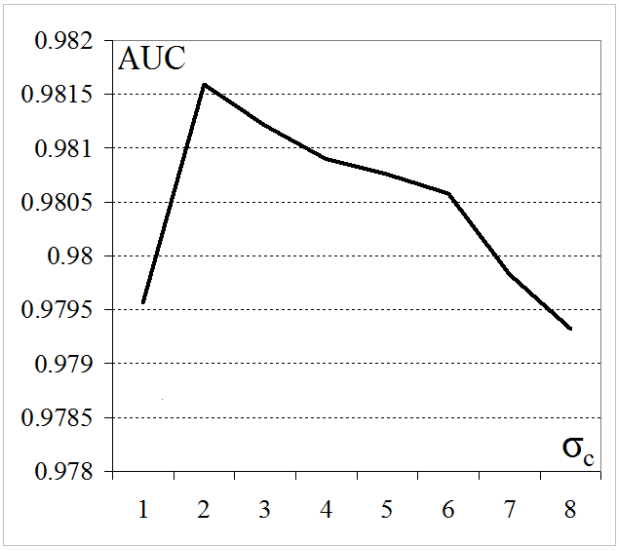

Рис. 4.8. График зависимости площади под ROC-кривой AUC от величины σ_c при зафиксированном значении $d = 21$.

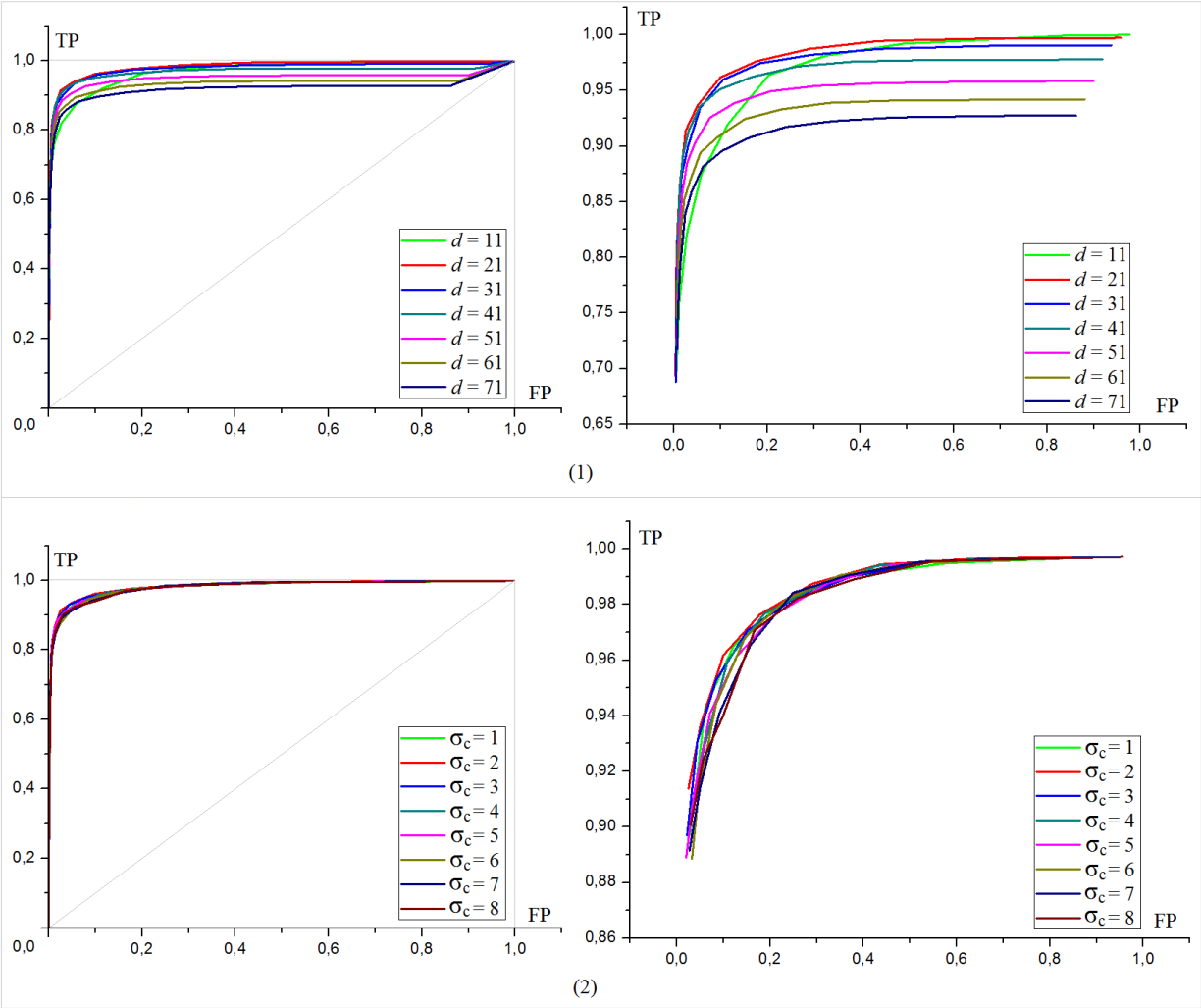

Рис. 4.9. Графики ROC-кривых для алгоритма, основанного на регуляризованном морфологическом проекторе: (1) — при изменении параметра d и оптимальном (по AUC) параметре σ_c, (2) — при изменении параметра σ_c и зафиксированном значении $d = 21$.

Слева — полный график, справа — увеличенный фрагмент в месте перегиба функции.

С ростом d площадь под графиками начинает равномерно уменьшаться. Следовательно, можно заключить, что размер сканирующего окна и для этого алгоритма зависит, в первую очередь, от размера объектов на изображениях. Также на рис. 4.9 (2) представлены ROC-кривые для различных значений σ_c (при фиксированном $d = 21$). Графики получились очень близки, но все же можно заметить, что наибольшую площадь имеет график, соответствующий $\sigma_c = 2$ (что подтверждает рис. 4.8).

Глобальная оптимизация.

Алгоритм поиска структурных различий, основанный на глобальной оптимизации специ-

альной энергетической функции, существенно отличается от всех вышеперечисленных своим набором параметров. Схема испытаний для него будет выглядеть следующим образом:

1. Генерация изображений с добавлением шума, стандартное отклонение которого выбирается из интервала $\sigma \in [0, 25]$ с шагом 0.5.

2. Выбор размера окрестности текущего пикселя $d \in [3, 15]$ с шагом 2, параметра, характеризующего структуру первого изображения $\alpha \in [3, 27]$ с шагом 2, и параметра близости приведения изображений друг к другу $\lambda \in [100, 2000]$ с шагом 100 (интервал варьирования параметров установлен на основе опыта работы с алгоритмом так, чтобы гарантированно содержать оптимальные значения).

3. К полученной паре изображений применяется алгоритм поиска структурных различий на основе глобальной оптимизации с выбранными параметрами d, α и λ. Результат работы алгоритма — результирующее разностное изображение R — после применения пороговой обработки сравнивается с ручной разметкой для нахождения значений TP, TN, FN, FP, и вычисляется величина AUC. Также определяется оптимальное значение порога T_{opt}, дающее наименьшее значение суммы ошибок первого и второго рода.

4. Шаги 1–3 повторяются 7 раз с одинаковыми значениями параметров, и полученные результаты усредняются.

5. Перебирая все значения параметров, формируются таблицы результатов $\mathrm{AUC}(\sigma, d, \alpha, \lambda)$ и $T_{opt}(\sigma, d, \alpha, \lambda)$, после чего определяются оптимальные значения параметров. Для зафиксированных d, α выбирается величина λ, дающая максимальное значение AUC (отдельно для каждого уровня шума σ). Далее для зафиксированного d определяется оптимальное значение α, после чего остаются таблицы результатов того же вида, что и для предыдущих алгоритмов: $\mathrm{AUC}(\sigma, d)$ и $T_{opt}(\sigma, d)$. Уже по этим сокращенным таблицам также строятся соответствующие графики, и для оптимального порога находится аппроксимирующая функция. Аппроксимация проводится с помощью метода наименьших квадратов полиномом пятой степени по переменной d и третей степени по σ (4.2). Для аппроксимации используется пакет OriginPro 8.0.

6. Фиксируя значение σ, определяются значения d, соответствующие максимальному значению AUC. Полученные данные используются для построения эмпирической зависимости размера сканирующего окна d от величины σ. Аппроксимация проводится линейными функциями с помощью пакета OriginPro 8.0.

Получены следующие результаты:

1. Выбор оптимальной величины параметра близости приведения исходных изображений друг к другу λ осуществлялся следующим образом: фиксировались d и α, и для каждого σ выбиралось то значение λ, при котором величина AUC была максимальной. Затем среди полученного

Таблица 4.3

Оптимальные значения параметра λ в зависимости от d и α

$d \setminus \alpha$	3	5	7	9	11	13	15	17	19	21	23	25	27
3	–	1900	2000	2000	1900	1700	1800	1800	1800	1700	1900	1800	1600
5	–	1900	1900	2000	2000	1900	2000	1600	1900	2000	1600	1600	1600
7	2000	2000	1700	1900	2000	1700	1600	1700	2000	1700	2000	1900	1700
9	2000	2000	1900	1700	1500	1500	1100	–	–	–	–	–	–
11	2000	2000	1500	2000	1700	1500	2000	–	–	–	–	–	–
13	2000	1800	1800	–	–	–	–	–	–	–	–	–	–
15	2000	1900	–	–	–	–	–	–	–	–	–	–	–

51-го значения (поскольку $\sigma \in [0, 25]$ с шагом 0.5) выбиралась мода. Перебирая значения d и α, была построена таблица 4.3, по которой можно судить о том, что никакой закономерности в полученных значениях не обнаруживается, но можно сделать замечание: оптимальные значения всегда попадают в интервал $[1100, 2000]$, и при больших α разброс значений становится больше. По всей видимости это связано с тем, что алгоритм, имея возможность приводить изображения друг к другу с произвольной точностью, аккуратно обрабатывал и добавленный случайный шум. Следовательно, разброс значений был вызван сильной зависимостью работы алгоритма от входных данных. Поэтому в качестве оптимального значения параметра выбрано простое среднее (по табл. 4.3): $\lambda = 1816$.

2. Параметр α, характеризующий структуру первого изображения, показал более регулярное поведение. На рис. 4.11 представлены графики зависимости величины $\alpha(\sigma)$ при различных значениях параметра d и оптимальной величине λ. Можно видеть, что при малых значениях d полученные графики аппроксимируются линейными функциями, а при больших — прямыми $\alpha = \text{const}$. Были получены следующие формулы:

$$\begin{gathered}\alpha = 0.39457 \cdot \sigma + 9.20513 \text{ при } d = 3,\\ \alpha = 0.27439 \cdot \sigma + 4.35445 \text{ при } d = 5,\\ \alpha = 0.14335 \cdot \sigma + 4.26697 \text{ при } d = 7,\\ \alpha = 5 \text{ при } d = 9,\ 11,\ 13,\ 15.\end{gathered} \tag{4.7}$$

Точность аппроксимации: $K_D(\alpha; \sigma, d = 3) = 0.90727$, $K_D(\alpha; \sigma, d = 5) = 0.92592$, $K_D(\alpha; \sigma, d = 7) = 0.75866$. Аппроксимировать графики при больших значениях d, очевидно, не имело смысла. На рис. 4.12 представлены графики ROC-кривых при различных значениях параметра α (уровень

Таблица 4.4

Коэффициенты аппроксимации оптимального порога алгоритма на основе глобальной оптимизации для формулы (4.2)

Коэффициент	z_0	a_1	a_2	a_3	a_4	a_5
Значение	6.86324	-4.80658	1.37414	$-1.87 \cdot 10^{-1}$	$1.17 \cdot 10^{-2}$	$-2.73 \cdot 10^{-4}$
Коэффициент	b_1	b_2	b_3	c_1	c_2	c_3
Значение	9.05255	-0.61693	0.01149	-4.11141	$9.95 \cdot 10^{-1}$	$-1.14 \cdot 10^{-1}$
Коэффициент	c_4	c_5	e_1	e_2	e_3	e_4
Значение	$6.12 \cdot 10^{-3}$	$-1.24 \cdot 10^{-4}$	0.33225	$-8.48 \cdot 10^{-2}$	$1.04 \cdot 10^{-2}$	$-5.99 \cdot 10^{-4}$
Коэффициент	e_5	h_1	h_2	h_3	h_4	h_5
Значение	$1.29 \cdot 10^{-5}$	$-6.59 \cdot 10^{-3}$	$1.82 \cdot 10^{-3}$	$-2.39 \cdot 10^{-4}$	$1.44 \cdot 10^{-5}$	$-3.21 \cdot 10^{-7}$

шума $\sigma = 10$, $d = 15$, $\lambda = 1816$). Наибольшую площадь имеет график, соответствующий $\alpha = 5$; близок к нему график при $\alpha = 3$. Эти значения можно считать оптимальными для данного параметра.

3. График зависимости $\mathrm{AUC}(\sigma, d)$ представлен на рис. 4.10. Можно провести те же рассуждения, что и для предыдущих алгоритмов: аппроксимировать зависимость величины оптимального порога T_{opt} от σ и d по формуле (4.2). Полученные коэффициенты аппроксимирующей функции — в таблице 4.4, точность аппроксимации $K_D(T_{opt}; d, \sigma) = 0.98541$.

На рисунке 4.13 представлены графики зависимости $\mathrm{AUC}(\sigma)$ для различных значений параметра размера окрестности пикселя d при оптимальном выборе α и λ. Можно видеть, что оптимальное значение $d = 15$ оказалось независящим от уровня шума. Также следует отметить увеличение значения AUC с ростом d: как было сказано ранее, для данного алгоритма большой размер окрестности играет роль сглаживания, что положительно влияет на результатах его работы.

В следующем пункте будут приведены результаты сравнения работы алгоритмов для модельных и реальных изображений с применением полученных соотношений для их оптимальных параметров.

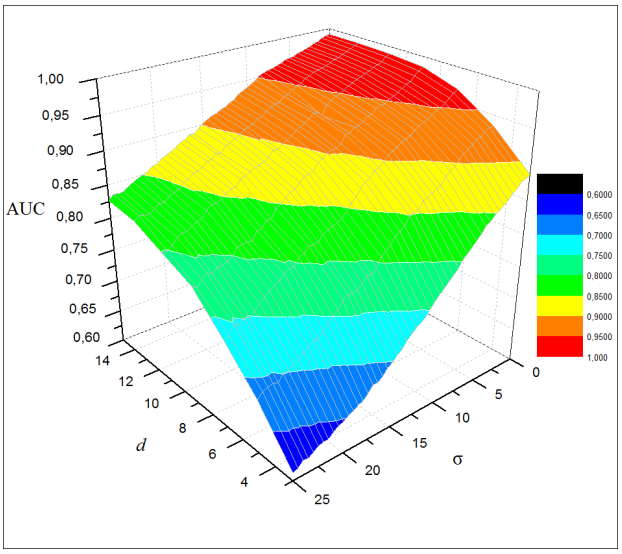

Рис. 4.10. Эмпирически полученный график зависимости площади под ROC-кривой AUC от размера окна d и стандартного отклонения шума σ для алгоритма, основанного на глобальной оптимизации

(для оптимальных значений α и λ).

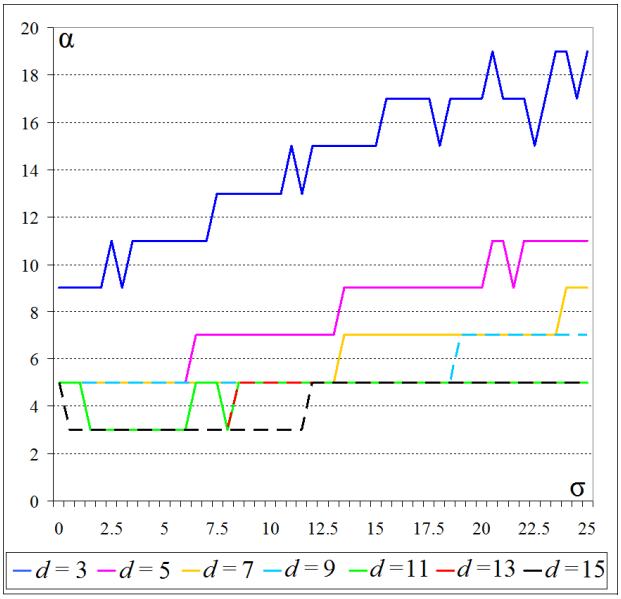

Рис. 4.11. Графики зависимости величины $\alpha(\sigma)$ для алгоритма, основанного на глобальной оптимизации, при различных значениях параметра d и оптимальной величине λ.

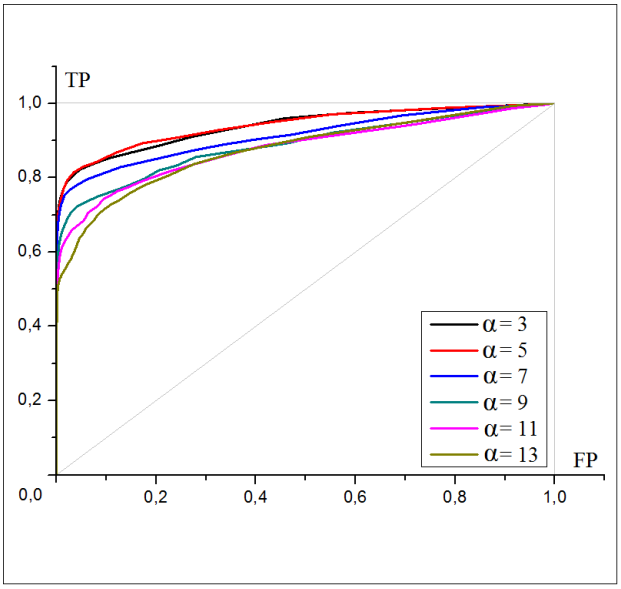

Рис. 4.12. Графики ROC-кривых для алгоритма, основанного на глобальной оптимизации при различных значениях параметра α

(уровень шума $\sigma = 10$, $d = 15$, $\lambda = 1816$).

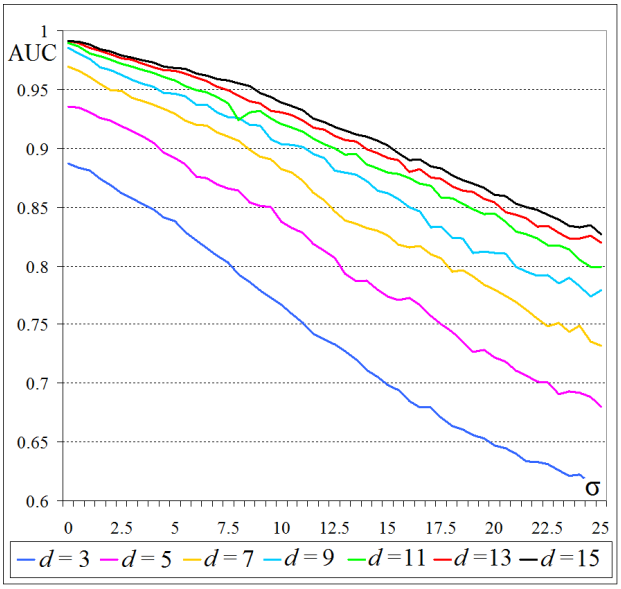

Рис. 4.13. Графики зависимости AUC(σ) при различных значений параметра d для алгоритма поиска структурных различий, основанного на глобальной оптимизации

(для оптимальных значений α и λ).

4.5 Сравнение результатов работы алгоритмов поиска структурных различий

Получив результаты исследования оптимальных параметров алгоритмов поиска структурных различий изображений, можно перейти к сравнению алгоритмов между собой с целью определения наиболее подходящего для создания программного комплекса. Сравнение будет проходить в два этапа: сначала будут рассмотрены результаты численного эксперимента из предыдущего пункта, затем алгоритмы будут применены к реальным снимкам.

На рис. 4.14 приведены результаты работы алгоритмов для созданных по модели из п. 4.2 изображений (приведенных на рис. 4.1), к которым добавлен шум с параметрами $N(0, 14)$. Черным цветом отмечены верно классифицированные положительные примеры, белым — верно классифицированные отрицательные случаи, красным — пропуск (ошибка I рода), зеленым — ложная тревога (ошибка II рода). Можно видеть, что:

— морфологический проектор выдал много мелкого шума с неплохой локализацией объектов, но пропустил одно изменение;

— линейная и квадратичная функции оказались близки — ложных различий не обнаружили, но с локализацией истинных не справились;

— регуляризованный морфологический проектор достаточно надежно обнаружил все имеющиеся различия с некоторыми ложными выбросами вдоль границ;

— метод, основанный на глобальной оптимизации, нашел все различия, но и выдал наибольшее количество ложных тревог.

Таблицы результатов эксперимента, содержащие площади под ROC-кривыми в зависимости от размера сканирующего окна и уровня шума, позволяют сравнить качество работы алгоритмов, основанных на морфологическом проекторе и его регуляризованном варианте, линейных и квадратичных функциях преобразования яркости. Для этого, зафиксировав значение d, можно поточечно сравнить полученные графики. Однако в алгоритме, основанном на глобальной оптимизации, параметр d обозначает размер окрестности пикселя и не является тождественным размеру сканирующего окна, из-за чего прямое их сопоставление становится некорректным. Поэтому от данного алгоритма для сравнения будут использоваться два графика: с параметром d, дающим наилучший и наихудший результаты, которые будут задавать границы качества работы алгоритма. Эти графики будут взяты из рис. 4.13 (от $d = 3$ до $d = 15$).

На рис. 4.15 приведены графики зависимости площади под ROC-кривой AUC от величины стандартного отклонения шума σ при различных значениях параметра d для алгоритмов поиска структурных различий изображений. При небольшом значении d наилучшее качество

работы демонстрирует алгоритм, основанный на квадратичных функциях преобразования яркости, но к значению $d = 41$ по качеству работы его превосходит алгоритм, основанный на регуляризованном варианте проектора, и в дальнейшем преимущество остается за этим алгоритмом. Что касается алгоритма, основанного на глобальной оптимизации, то в лучшем случае он превосходит остальные алгоритмы только при больших значениях d, а в худшем — уступает всем. Таким образом, наилучшее качество работы показали алгоритмы, основанные на квадратичных функциях и регуляризованном морфологическом проекторе. Однако, такое сравнение несколько неоднозначно: для фиксированного значения d при различных уровнях шума наибольшее значение AUC может соответствовать различным алгоритмам. Поэтому предлагается построить еще два графика — для среднего и максимального значения AUC.

Среднее значение по уровню шума σ величины площади под ROC-кривой AUC вычисляется по зафиксированному размеру сканирующего окна d для всех алгоритмов, кроме основанного на глобальной оптимизации, т. е. это среднее качество работы алгоритма для зафиксированного набора параметров при различном уровне шума. Полученные графики представлены на рис. 4.16. Можно видеть, что до $d = 29$ лучшим является алгоритм, основанный на квадратичных функциях преобразования яркости, а после — на регуляризованном проекторе.

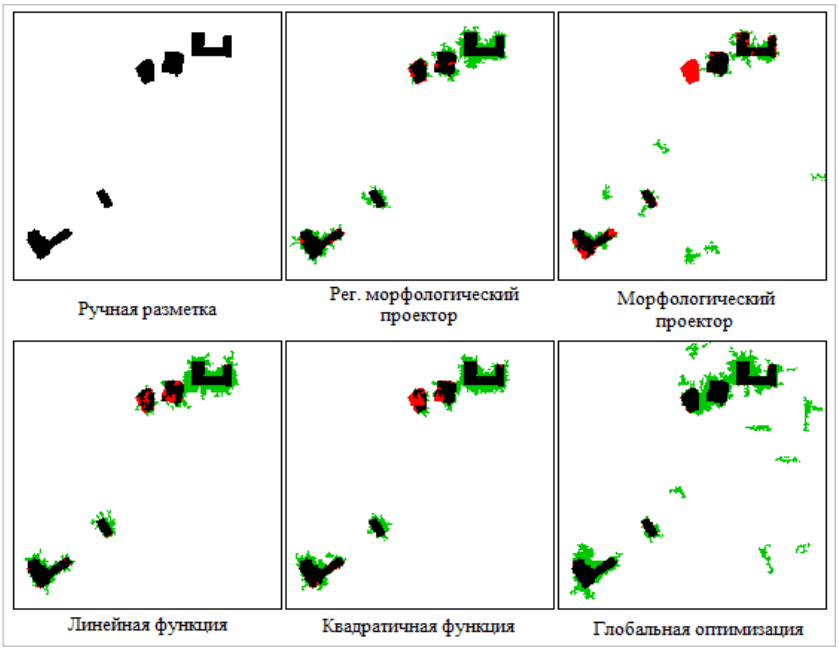

Рис. 4.14. Ручная разметка и результаты работы алгоритмов поиска структурных различий с оптимальными параметрами для представленных на рис. 4.1 синтезированных изображений: черным цветом отмечено верное обнаружение, белым — верно классифицированные отрицательные случаи, красным — пропуск, зеленым — ложная тревога.

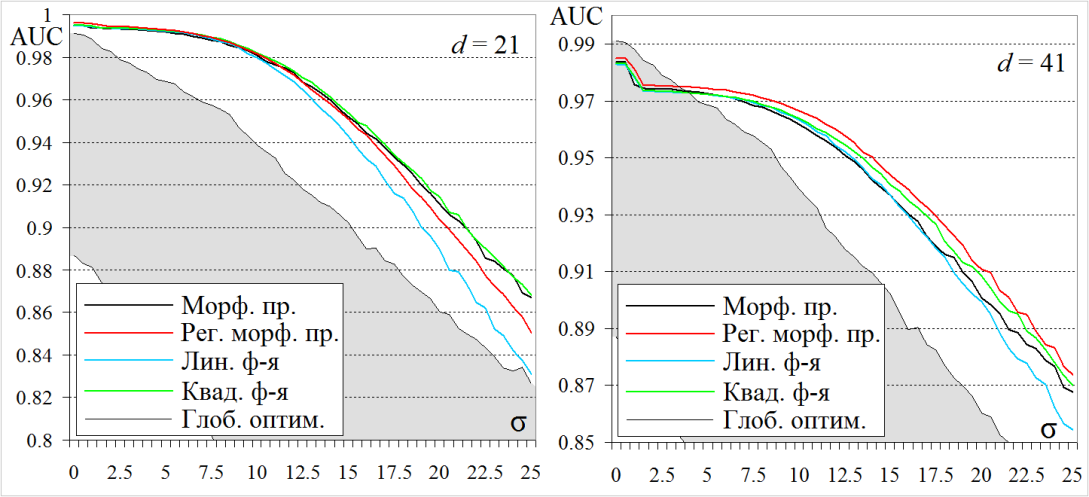

Рис. 4.15. Графики зависимости площади под ROC-кривой AUC от величины стандартного отклонения шума σ при различных значениях параметра d для алгоритмов поиска структурных различий изображений. Заштрихованная область соответствует графикам AUC(σ) алгоритма, основанного на глобальной оптимизации от $d = 3$ (нижняя граница) до $d = 15$ (верхняя граница).

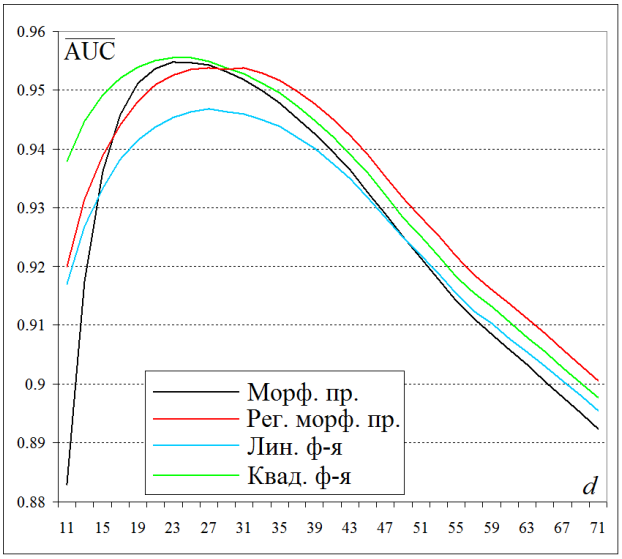

Рис. 4.16. Графики зависимости средней по уровню шума величины площади под ROC-кривой AUC от величины размера сканирующего окна d для алгоритмов поиска структурных различий изображений.

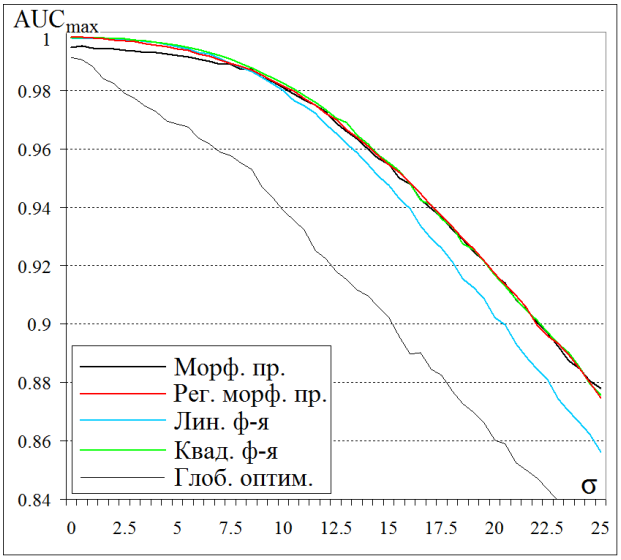

Рис. 4.17. Графики зависимости максимальной по всем параметрам величины площади под ROC-кривой AUC от величины стандартного отклонения шума σ для алгоритмов поиска структурных различий изображений.

Максимальное значение AUC находится для каждого значения σ по всем параметрам алгоритмов, т. е. это наилучшее качество работы алгоритма при заданном уровне шума. Полученные графики изображены на рис. 4.17. Снова наилучшие результаты оказались у алгоритмов, основанных на квадратичных функциях и регуляризованном проекторе: их графики практически совпадают. Заметим, что алгоритм, основанный на глобальной оптимизации, оказался наихудшим. Таким образом, по результатам численного эксперимента можно заключить, что наилучшее качество работы показали алгоритмы, основанные на регуляризованном морфологическом проекторе и на квадратичных функциях преобразования яркости. Второй имеет преимущество при небольших значениях параметра d. Однако, как правило, уровень шума на реальных снимках примерно $\sigma \in [14, 24]$, следовательно по формулам 4.3–4.6 оптимальное значение $d \in [23, 35]$. Этот интервал соответствует области, где алгоритм с регуляризованным проектором имеет преимущество, поэтому он и признается наилучшим по результатам численного эксперимента.

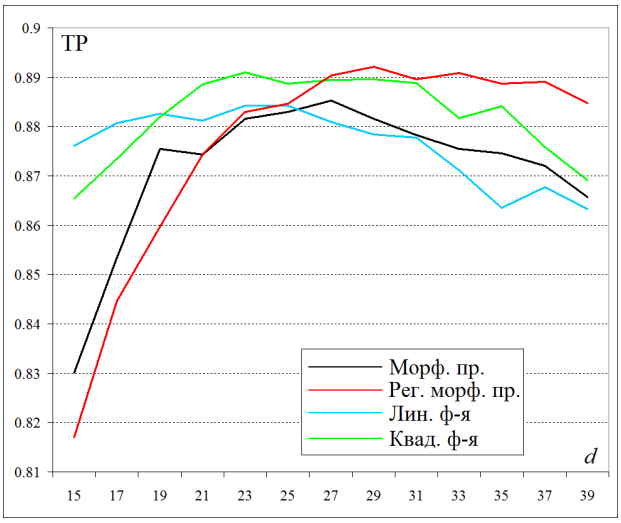

Рис. 4.18. Графики наибольшего уровня верного обнаружения TP алгоритмов для заданного максимально допустимого уровня ложной тревоги $FP = 0.1$ при различных значениях размера сканирующего окна d.

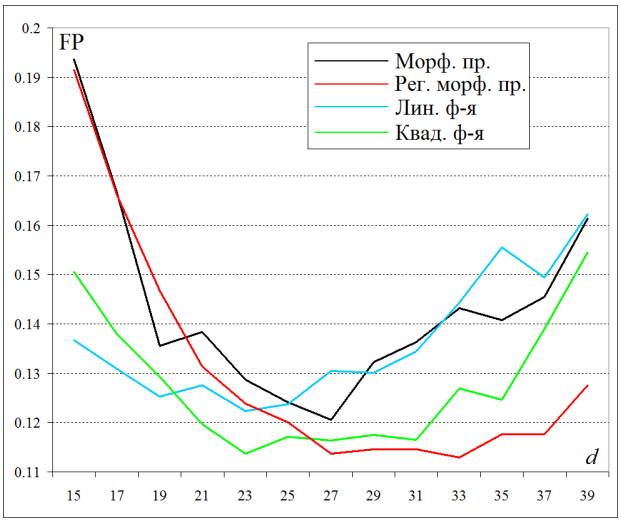

Рис. 4.19. Графики наименьшего уровня ложной тревоги FP алгоритмов для заданного минимально допустимого уровня верного обнаружения $TP = 0.9$ при различных значениях размера сканирующего окна d.

В реальной задаче может потребоваться знание следующих показателей: какой наилучший процент верного обнаружения (ложной тревоги) можно получить при зафиксированном максимально допустимом уровне ложной тревоги (минимально допустимом уровне верного обнаружения)? Для этого зададим допустимый уровень ложной тревоги $FP = 0.1$ и для каждого алгоритма определим наибольшее значение верного обнаружения TP. Полученные графики

Таблица 4.5

Максимальные значения TP (минимальные FP) при зафиксированном уровне $FP = 0.1$ ($TP = 0.9$) для модели из п. 4.2 работы алгоритмов, и оптимальные параметры, при которых они получены

Алгоритм	TP_{max}	FP_{min}	**Параметры**
Рег. морфологический проектор	0.892091	0.113738	$d = 29$, $\sigma_c = 2$
Квадратичная функция	0.890949	0.113725	$d = 23$
Морфологический проектор	0.885229	0.120545	$d = 27$
Линейная функция	0.884299	0.122274	$d = 23$
Глобальная оптимизация	0.80634	0.332012	$d = 19$, $\lambda = 1816$, $\alpha = 5$

в зависимости от размера сканирующего окна d представлены на рис. 4.18. Аналогично, на рис. 4.19 представлены графики наименьшего уровня ложной тревоги FP алгоритмов для заданного уровня верного обнаружения $TP = 0.9$. В обоих случаях можно видеть, что до значения $d = 19$ лучшие результаты показывает алгоритм, основанный на линейной функции преобразования яркости, от $d = 21$ до $d = 25$ — алгоритм, основанный на квадратичной функции преобразования яркости, а далее — алгоритм, основанный на регуляризованном морфологическом проекторе, что еще раз подтверждает полученные выше результаты. В таблице 4.5 представлены максимальные значения верного обнаружения TP (ложной тревоги FP) для всех алгоритмов и оптимальные параметры, при которых они получены.

Второй этап сравнения — проверка на реальных изображениях. Для этого были взяты два полутоновых снимка одного и того же участка городской застройки, сделанные с разницей в 4 года (рис. 4.20). Левый снимок будем называть первым (именно он был использован для создания синтезированных изображений в численном эксперименте), правый — вторым. Присутствующий на изображениях шум имеет оценку $\sigma = 14.5$. К этим снимкам применялись разработанные алгоритмы поиска структурных различий с оптимальными параметрами, определенными по формулам из предыдущего пункта и собранными в таблице 4.6.

На рис. 4.21 представлены фрагменты результатов радиометрической коррекции, приводящей яркости первого изображения ко второму с сохранением структуры, а также соответствующий фрагмент самого второго изображения, из которого будет вычитаться преобразованное. Можно видеть, что все методы успешно сопоставляют яркости двух изображений, но при этом морфологический проектор и глобальная оптимизация дают несколько зашумленный результат, линейная и квадратичная функция, наоборот, — размытый, а вот регуляризованный проектор оказывается золотой серединой, выдав гладкое изображение без потери четкости границ.

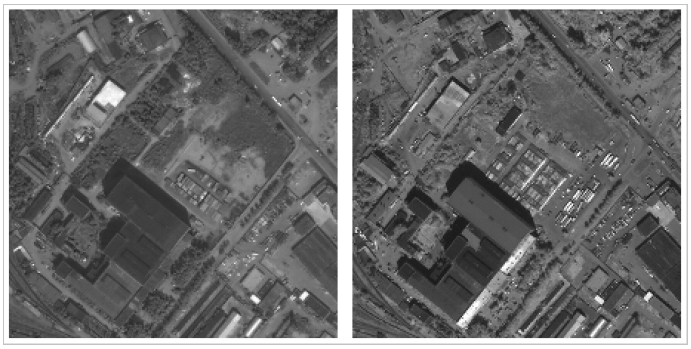

Рис. 4.20. Исходные полутоновые снимки городской застройки, сделанные с разницей в 4 года.

Таблица 4.6

Оптимальные параметры алгоритмов поиска структурных различий изображений, вероятности верного обнаружения и ложной тревоги, а также площади под ROC-кривыми

Алгоритм	d	**Другие параметры**	T_{opt}	**TP**	**FP**	**AUC**
Рег. морфологический проектор	23	$\sigma_c = 2$	20	73.1%	2.1%	0.897085
Квадратичная функция	23	–	22	68.7%	5.9%	0.896573
Морфологический проектор	23	–	21	66.9%	6.3%	0.895371
Линейная функция	25	–	22	63.3%	4.3%	0.892668
Глобальная оптимизация	15	$\lambda = 1816$, $\alpha = 5$	22	63.5%	10.6%	0.874071

Наконец, на рис. 4.22 изображены фрагменты созданного вручную эталона — ручная разметка присутствующих на снимках структурных различий и результатов работы алгоритмов. Черным цветом отмечены верно классифицированные положительные примеры, белым — верно классифицированные отрицательные случаи, красным — пропуск (ошибка I рода), зеленым — ложная тревога (ошибка II рода). Результаты таковы:

— морфологический проектор достаточно неплохо нашел и локализовал имеющиеся структурные различия, но и выдал большое количество ложных срабатываний;

— линейная и квадратичная функции оказались близки — среднее количество найденных различий и умеренное ложных тревог;

— регуляризованный морфологический проектор нашел практически те же различия, что и предыдущие алгоритмы, но зато выдал наименьшее количество ложных срабатываний;

— глобальная оптимизация показала худший результат по обнаружению и локализации различий, и уровень ложных тревог оказался наиболее значительным.

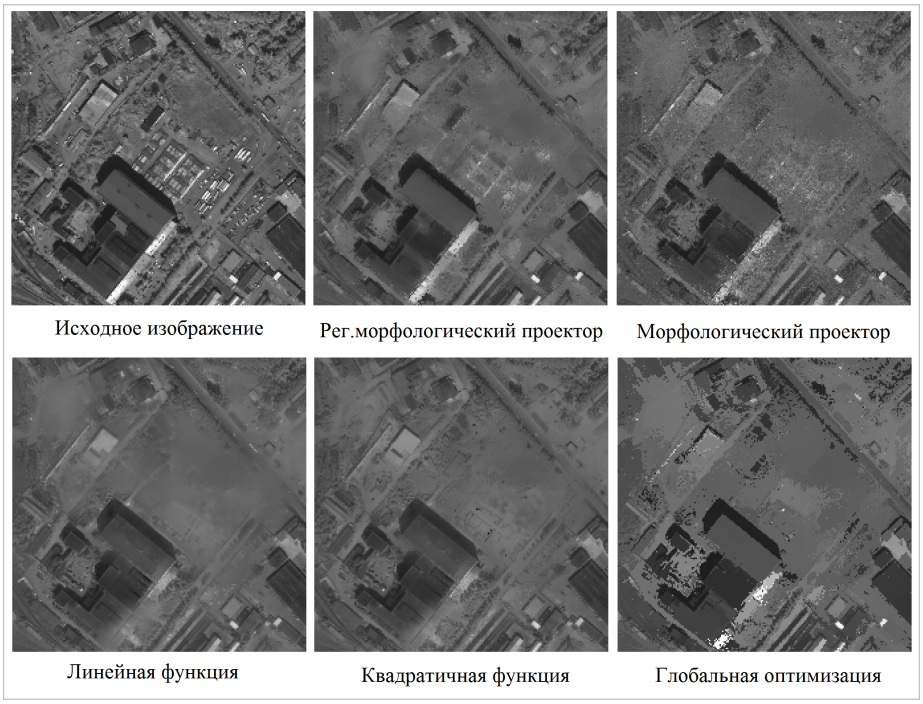

Рис. 4.21. Один из двух исходных снимков, к которому приводилось по яркости другое изображение, и результаты радиометрической коррекции, выполненной алгоритмами поиска структурных различий.

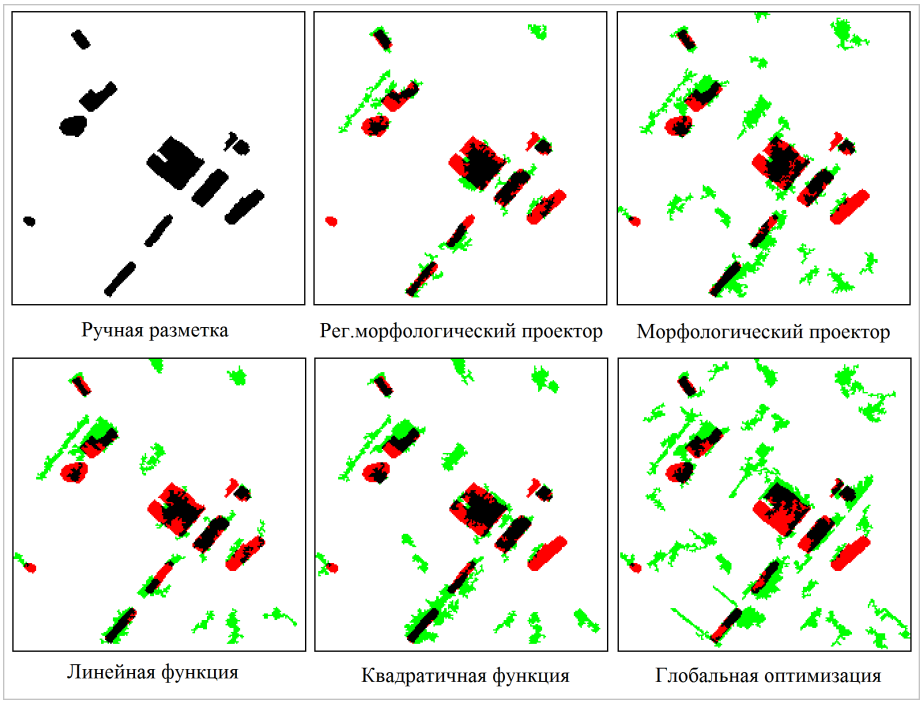

Рис. 4.22. Ручная разметка и результаты работы алгоритмов поиска структурных различий с оптимальными параметрами: черным цветом отмечено верное обнаружение, белым — верно классифицированные отрицательные случаи, красным — пропуск, зеленым — ложная тревога.

Площади под ROC-кривыми представлены в табл. 4.6, и наибольшая из них соответствует алгоритму, основанному на регуляризованном варианте морфологического проектора, что соответствует результатам, полученным в численном эксперименте.

4.6 Исследование влияния рассинхронизации изображений на поиск структурных различий

Один из важнейших аспектов работы алгоритма — это его устойчивость к качеству входных данных. До сих пор исследование проводилось относительно влияния случайного шума, присутствующего на снимках. Однако, важное значение имеет геометрическая выравненность изображений — как было указано во введении, она заключается в том, что каждый пиксель на обоих снимках имеет одни и те же координаты на местности, т. е. эти снимки одинаково привязаны к местности. Нарушение этого условия, очевидно, приводит к появлению ложных срабатываний алгоритмов в окрестностях объектов. Исследованию влияния подобного несоответствия посвящены работы [58, 117], в которых приведено описание эксперимента по оценке влияния рассинхронизации для изображения и его копии, благодаря чему исключено влияние структурных различий, и та же оценка для рассинхронизированных разновременных изображений (подробнее см. п. 1.2.1).

В данном пункте приведено исследование влияния рассинхронизации входных изображений на качество работы алгоритмов, описанных в главе 3, на основе подхода, аналогичного используемому в [117], но для более существенных сдвигов в несколько пикселей: подобное исследование интересно с точки зрения несовершенства (или отсутствия) геопривязки входных снимков или, что более вероятно, разного угла съемки. Для этого будет взят космический снимок земной поверхности (городская застройка), а в качестве второго изображения будет рассмотрена копия первого, сдвинутого на величину sd. Полученная пара изображений, разумеется, не содержит никаких структурных различий, но из–за отсутствия геометрической выровненности таковыми станут сместившиеся границы объектов и т.д. Следовательно, относительная площадь обнаруженных различий будет эквивалентна вероятности ложной тревоги. На рис. 4.23 представлены полученные графики зависимости ошибки от величины сдвига изображений относительно друг друга.

Отложенные на горизонтальной оси значения sd не целые, так как сдвиг осуществляется на целое значение по горизонтали и вертикали, а sd — это расстояние между пикселями после сдвига. Можно видеть, что наиболее устойчивым к возмущению входных данных оказался алгоритм, основанный на регуляризованном варианте морфологического проектора. Далее идет

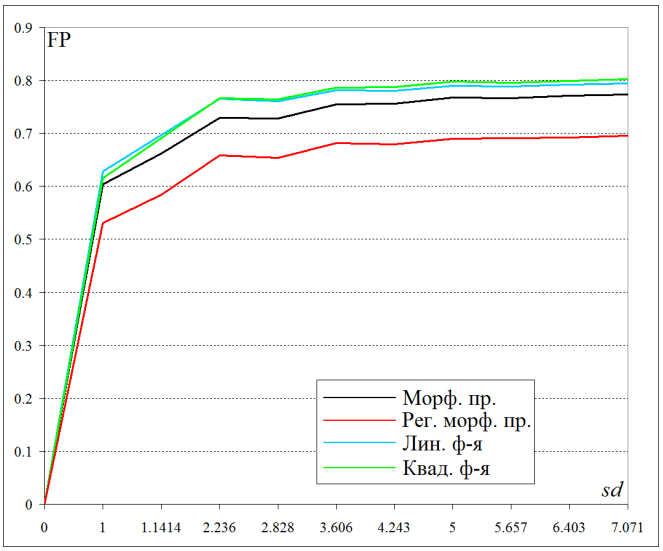

Рис. 4.23. Графики зависимости уровня ложной тревоги (false positive, FP) от величины рассинхронизации изображений sd, изначально не содержащих структурных различий, для различных алгоритмов.

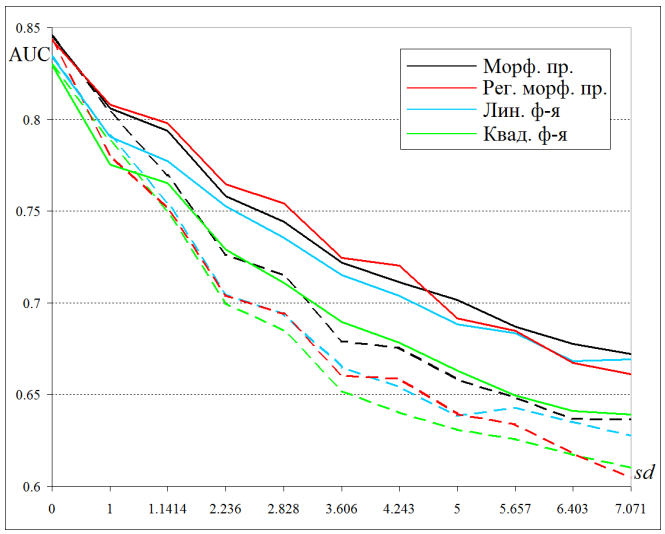

Рис. 4.24. Графики зависимости AUC(sd) для различных алгоритмов поиска структурных различий изображений без геометрической коррекции (пунктирные линии) и с ее использованием (сплошные линии). Размер сканирующего окна для всех алгоритмов $d = 21$.

алгоритм, основанный на морфологическом проекторе, а линейные и квадратичные функции продемонстрировали практически идентичные результаты.

В п. 3.5 было приведено описание метода, применимого для локальной геометрической коррекции входных снимков. Для проверки этого подхода, возьмем два геометрически выровненных космоснимка городской застройки и добавим в них локальный сдвиг на величину sd. После чего применим к полученным изображениям описанные выше алгоритмы поиска структурных различий — без использования “степени биективности” и с ней. Полученные результаты сравним с ручной разметкой; критерием качества работы по прежнему будет анализ ROC–кривых.

На рис. 4.24 изображены графики зависимости значения AUC от величины сдвига для алгоритмов без геометрической коррекции (пунктирные линии) и с ее использованием (сплошные линии). Во всех случаях применение “степени биективности” позволило повысить качество работы.

4.7 Выводы по главе 4

Проводя параллели между результатами работы на синтезированных и реальных снимках, можно сделать следующие выводы:

1. Алгоритм, основанный на морфологическом проекторе, неплохо находит и локализует структурные различия, но при этом выдает наибольшее количество мелких ложных срабатываний, что связано с тем, что присутствующий на изображениях шум сильно влияет на его работу.

2. Алгоритмы, основанные на линейных и квадратичных функциях, уступают морфлогическому проектору в количестве и качестве найденных различий, зато выдают меньшее число ложных тревог, что обусловлено самой сутью их определения: ограничение класса допустимых преобразований позволяет приводить изображения друг а другу более грубо, за счет чего происходит игнорирование большей части случайного шума.

3. Алгоритм, основанный на регуляризованном варианте морфологического проектора, не уступая предыдущим методам в качестве обнаружения имеющихся структурных различий, выдает заметно меньшее количество ложных срабатываний. Также он является наиболее устойчивым к возможной рассинхронизации входных изображений.

4. Алгоритм, основанный на глобальной оптимизации, показал наиболее неоднозначные результаты: качественное обнаружение и локализация различий, но и большое количество ложных тревог, что говорит о его крайней чувствительности к входным данным.

В то же время алгоритм, основанный на глобальной оптимизации, имеет наиболее высокую сложность настройки (из-за большего числа параметров и связанной со случайным выбором работой) и наибольшую длительность работы (см. п. 3.8). К тому же, если бы исходные снимки имели лучшее качество, то преимущество могло остаться за другим алгоритмом. Лучшим среди оставшихся является алгоритм, основанный на регуляризованном проекторе, и он же был лучшим в численном эксперименте. Поэтому для создания программного комплекса, описанного в предыдущей главе, был выбран алгоритм, основанный на регуляризованном морфологическом проекторе. Метод, основанный на глобальной оптимизации, является крайне перспективным, но все же требует доработки.

Заключение

В представленной работе была рассмотрена задача поиска структурных различий изображений по двум разновременным цифровым космическим снимкам земной поверхности. Поиск структурных различий используется во многих приложениях и областях компьютерного и технического зрения, в частности, при обновлении топографических карт, где он играет важнейшее значение. Обладая столь значительной практической ценностью, поиск различий, очевидно, требует качественного исследования, и предложенные в данной диссертации алгоритмы позволяют успешно решать поставленную задачу.

Поиск различий имеет столь широкое применение благодаря вариативности понятия “различие”: в разных задачах это может быть изменение формы, цвета и т.д., или же изменения определенного вида (например, появление заданного типа объектов). Соответственно, в данной работе сначала была введена формализация понятий структуры изображения и структурных различий как изменений типа появление / исчезновение объектов на снимках. Была указана связь диссертации с существующей системой методов. Далее была предложена математическая модель структурных различий пары изображений, и на основе этой модели сформулирована задача поиска структурных различий.

Для поставленной задачи был разработан набор алгоритмов, решающих ее. Базовый способ поиска различий заключается в попиксельном вычитании одного изображения из другого. Однако, снимки, сделанные с интервалом в несколько лет, существенно различаются по яркости даже в тех областях, где отсутствуют различия. Поэтому предварительно необходимо провести выравнивание яркостей изображений — в литературе это называется относительной радиометрической коррекцией входных снимков. В качестве методов проведения такой коррекции были предложены алгоритмы, основанные на морфологическом проекторе и его регуляризованном варианте, линейных и квадратичных функциях преобразования яркости и глобальной оптимизации. Также был предложен вариант морфологического проектора, работающий с цветными изображениями, и метод, ищущий структурные различия без использования радиометрической коррекции входных снимков.

Качество работы каждого алгоритма напрямую зависит от выбранного набора его па-

раметров, среди которых можно выделить следующие: общие — оптимальный порог и размер сканирующего окна (размер окрестности пикселя для алгоритма, основанного на глобальной оптимизации); частные (для алгоритмов, основанных на регуляризованном морфологическом проекторе и глобальной оптимизации). Как и во многих других задачах обработки изображений, ключевую роль играет оптимальный порог. Критерием оптимальности был выбран минимум суммы ошибок первого и второго рода. Для алгоритма, основанного на морфологическом проекторе, для определенной ранее математической модели была получена формула распределения вероятностей яркости зафиксированного пикселя изображения, представляющего собой разность исходного изображения и его морфологической проекции. На основе этой формулы было получено выражение для оптимального порога данного алгоритма.

Определение всех остальных оптимальных параметров и сравнение результатов работы алгоритмов между собой потребовали проведения численного эксперимента. Вследствие сложности и, зачастую, неоднозначности обрабатываемых данных, была предложена еще одна математическая модель, более близкая к реальным данным, с помощью которой были определены оптимальные параметры каждого алгоритма. Сравнение методов между собой показало преимущество алгоритма, основанного на регуляризованном морфологическом проекторе, который был рекомендован для создания прикладного приложения. На основе всех полученных данных был разработан программный комплекс, решающий задачу поиска структурных различий изображений, который был интегрирован в среду визуализации и обработки данных дистанционного зондирования Земли ENVI.

На защиту выносятся следующие новые научные результаты диссертационной работы:

1. На основе понятия структуры изображения как набора множеств уровней яркости введена вероятностная модель структурных различий пары изображений и сформулирована задача поиска структурных различий. Для этой модели впервые получена конструктивная формула распределения вероятности яркости точек разностного изображения, представляющего собой разность исходного изображения и его морфологической проекции. С помощью этой формулы построен оптимальный порог исследуемого алгоритма.

2. Разработан ряд новых прикладных алгоритмов, решающих задачу поиска структурных различий изображений, использующих как сканирующее окно, так и глобальную оптимизацию энергетической функции. Также разработан алгоритм, основанный на непосредственном применении введенного определения структурных различий, полезный как для поиска малоразмерных структурных различий, так и для устранения локальной рассинхронизации изображений.

3. Предложена специализированная методика численного эксперимента для сравнения качества работы алгоритмов поиска структурных различий изображений. Эта методика построена

на базе параллельных суперкомпьютерных технологий и, в частности, ее программная реализация использует высокопроизводительные многоядерные графические процессоры. С помощью данной методики проведен объемный вычислительный эксперимент, результаты которого позволили оценить оптимальные параметры новых и существующих алгоритмов.

4. Созданный на базе проведенного исследования программный комплекс, решающий задачу оперативной оценки изменчивости топографических объектов местности по разновременным космическим снимкам. Данный комплекс был интегрирован в среду визуализации и обработки данных дистанционного зондирования Земли ENVI и внедрен в практику научно–исследовательского и производственного центра АО "НИиП центр Природа", (г. Москва).

Резюмируя, можно отметить, что в ходе исследовательской работы удалось получить формализацию и удовлетворительное решение задачи поиска структурных различий с помощью нескольких методик, но в тоже время на основе единого теоретического подхода. Для основных результатов диссертации можно сформулировать следующие направления дальнейшей работы.

1. Получить теоретические формулы для оптимального порога алгоритмов поиска структурных различий. В настоящий момент такая формула есть лишь для алгоритма, основанного на морфологическом проекторе. Определив эти формулы для остальных алгоритмов, можно будет сформулировать утверждение о теоретическом преимуществе одного из них. Эмпирическое преимущество осталось за алгоритмом, основанном на регуляризованном морфологическом проекторе.

2. Использование глобальной оптимизации дает очень неплохой и перспективный результат, поэтому важным направлением является усовершенствование данного метода и его дальнейшее применение в прикладных задачах.

3. Разработать версии алгоритмов, работающие с многоканальными изображениями (цветными, но не только: например, перспективен формат NDVI). На данный момент используются полутоновые космические снимки, но за счет цвета можно получить гораздо больше полезной информации, которая позволит более качественно находить структурные различия. Возможно, стоит использовать также данные различных геофизических полей. Например, поле высот строений и сооружений, которое можно получить, используя стереоснимки (такие работы в мире уже существуют, хотя и для аэросъемки).

Публикации автора по теме диссертации

- Публикации в изданиях, рекомендованных ВАК:

1. *Корнилов Ф.А.* Исследование алгоритма поиска структурных различий изображений // Труды Института математики и механики УрО РАН. Екатеринбург: ИММ УрО РАН, 2013. Т. 19, № 1. С. 144–159.

2. *Бровко Е.А., Кандоба И.Н., Корнилов Ф.А., Перевалов Д.С.* Оперативный мониторинг структурных изменений изображений объектов на космических снимках земной поверхности // Программные продукты и системы, 2015. № 1 (109). С. 79–86.

- Публикации в других изданиях:

3. *Корнилов Ф.А., Перевалов Д.С.* Задача обнаружения структурных различий изображений // Алгоритмы и програм. средства параллел. вычислений. Екатеринбург: УрО РАН, 2011. Вып. 11. С. 37–56.

4. *Корнилов Ф.А.* Поиск структурных различий изображений: алгоритмы и методика исследования // Машинное обучение и анализ данных, 2014. Т. 1. № 7. С.902–919.

5. *Корнилов Ф.А., Костоусов В.Б., Перевалов Д.С.* Быстрые алгоритмы анализа изображений со структурными искажениями // Труды научно-технической конференции-семинара "Техническое зрение в системах управления мобильными объектами-2010". М.: КДУ, 2011. С. 303–309.

6. *Корнилов Ф.А., Перевалов Д.С.* О применении регуляризации морфологического проектора в задаче поиска структурных различий // Сборник трудов научно-технической конференции "Техническое зрение в системах управления 2012". Москва, 2012. С. 265–271.

7. *Корнилов Ф.А.* Исследование влияния рассинхронизации входных изображений на качество работы алгоритмов поиска структурных различий // Машинное обучение и анализ данных, 2015. Т. 1, № 12. С.1686–1695.

8. *Корнилов Ф.А., Перевалов Д.С.* Обнаружение изменений объектов земной поверхности по спутниковым снимкам // Проблемы теоретической и прикладной математики. Тезисы 41-й Всероссийской молодежной конференции. Екатеринбург: ИММ УрО РАН, 2010. С. 534–540.

9. *Корнилов Ф.А., Перевалов Д.С.* Оценка оптимального порога для алгоритма поиска структурных изменений с помощью проектора Пытьева // Современные проблемы математики. Тезисы 42-й Всероссийской молодежной школы-конференции. Екатеринбург: ИММ УрО РАН, 2011. С. 291–293.

10. *Корнилов Ф.А.* Алгоритм поиска структурных различий на основе Марковских случайных полей // Современные проблемы математики. Тезисы Международной (43-й Всероссийской) молодежной школы-конференции. Екатеринбург: ИММ УрО РАН, 2012. С. 311–313.

11. *Корнилов Ф.А.* Теоретическое исследование морфологического проектора Пытьева // Современные проблемы математики. Тезисы Международной (44-й Всероссийской) молодежной школы-конференции. Екатеринбург: ИММ УрО РАН, 2013. С. 322–325.

12. *Корнилов Ф.А., Перевалов Д.С.* Алгоритм поиска структурных изменений на изображениях // СПИСОК-2009: материалы межвузовской научной конференции по проблемам информатики. Екатеринбург: Издательство Урал. ун-та, 2009. С. 264–266.

13. *Корнилов Ф.А., Перевалов Д.С.* Алгоритмы обнаружения структурных изменений на разновременных снимках // Тез. и аннот. докл. рег. науч.-тех. конф. “Данные дистанционного зондирования Земли для решения практических задач развития территорий”. Екатеринбург: ФГУП “Уралгеоинформ”, 2009. С. 21–22.

14. *Корнилов Ф.А.* Поиск структурных различий изображений // Интеллектуализация обработки информации: 10-я международная конференция. Греция, о. Крит, 6–10 октября 2014 г.: Тезисы докладов. — М.: Торус Пресс, 2014. С.118–119.

15. *Корнилов Ф.А.* Исследование влияния рассинхронизации входных изображений на качество работы алгоритмов поиска структурных различий // Математические методы распознавания образов: Тезисы докладов 17-й Всероссийской конференции с международным участием, г.Светлогорск, 2015 г. — М.: Торус Пресс, 2015. С.160–161.

В работе [4] автору принадлежит разработка структуры программного комплекса и реализация алгоритма поиска структурных различий, а его соавторам — формулировка технического задания и оформление документации. В работах [17–19, 21, 24] автору принадлежат основные результаты, а его соавтору — постановка задачи. В работе [20] автору принадлежит раздел 3 (стр. 306). В работе [22] автору принадлежат основные результаты, а его соавтору — постановка задачи и аппарат структурных функций (стр. 39).

Исследования, проведенные в диссертации, были поддержаны программами фундаментальных исследований Президиума РАН при финансовой поддержке УрО РАН “Динамические системы и теория управления” (проект 12-П-1-1022) и “Информационные, управляющие и интеллектуальные технологии и системы” (проект 12-П-1-1023).

Автор благодарен своему научному руководителю В. Б. Костоусову за постановку задачи и постоянную поддержку в работе, Д. С. Перевалову за ценное обсуждение результатов и полезные замечания, а также М. Ю. Хачаю и Г. Ф. Корниловой за неоценимую помощь в оформлении диссертации.

Литература

[1] **Айвазян С.А., Мхитарян В.С.** Прикладная статистика и основы эконометрики. М.: ЮНИТИ, 1998. 1005 с.

[2] **Бердышев В.И., Костоусов В.Б.** Экстремальные задачи и модели навигации по геофизическим полям. Екатеринбург: УрО РАН, 2007. 270 с.

[3] **Боресков А.В. и др.** Параллельные вычисления на GPU. Архитектурная и программная модель CUDA. М.: Изд-во Моск. ун-та. Сер. “Суперкомпьютерное образование”, 2012. 336 с.

[4] **Бровко Е.А., Кандоба И.Н., Корнилов Ф.А., Перевалов Д.С.** Оперативный мониторинг структурных изменений изображений объектов на космических снимках земной поверхности // Программные продукты и системы, 2015. № 1 (109). С. 79–86.

[5] **Визильтер Ю.В.** Обобщенная проективная морфология // Компьютерная оптика, 2008. Т. 32, № 4. С. 384–399.

[6] **Визильтер Ю.В., Горбацевич В.** Морфологическая сегментация изображений земной поверхности с использованием стековых деревьев // 20-я междунар. конф. по компьютерной графике и зрению. Санкт-Петербург, 2010. С. 175–180.

[7] **Визильтер Ю.В.** Теория и методы морфологического анализа изображений: дис. на соискание ученой степени д-ра физ.-мат. наук. Москва, 2008. 323 с.

[8] **Винклер Г.** Анализ изображений, случайные поля и динамические методы Монте-Карло. Математические основы. Новосибирск: Изд-во СО РАН, филиал “Гео”, 2002. 343 с.

[9] **Гасилов В.Л., Костоусов В.Б.** Задача идентификации движущегося объекта на основе обработки изображения внешнего информационного поля // Изв. РАН. Сер. Техническая кибернетика, 1994. № 3. С. 78–86.

[10] **Гасилов В.Л., Костоусов В.Б.** Методы получения и представления эталонной информации о геофизических полях // Материалы XX НТК памяти Н. Н. Острякова. СПб: Гироскопия и навигация, 1996. № 4. С. 64–65.

[11] **Гонсалес Р., Вудс Р.** Цифровая обработка изображений. Пер. с англ. М.: Техносфера, 2005. 1072 с.

[12] **Ефимов С.А., Кандоба И.Н., Костоусов В.Б., Перевалов Д.С., Скрипнюк В.В.** Система автоматизированного топографического мониторинга по снимкам земной поверхности // Геодезия и картография, 2008. № 12. С. 34–41.

[13] **Кандоба И.Н., Костоусов В.Б., Перевалов Д.С.** Создание математических методов, параллельных алгоритмов и программ для решения задач анализа изображений и задач управления в системах высокоточной навигации и наведения движущихся объектов // Докл. 13-й всеросс. конф. "Матем. методы распознавания образов" (ММРО–13). М.: МАКС Пресс, 2007. С. 335–337.

[14] **Кандоба И.Н., Костоусов В.Б., Костоусов К.В., Перевалов Д.С.** Алгоритмы поиска и классификации изображений линейных объектов на космоснимках // Докл. 14-й всеросс. конф. "Матем. методы распознавания образов" (ММРО–14). М.: МАКС Пресс, 2009. С. 536–539.

[15] **Кобзарь А.И.** Прикладная математическая статистика: для инженеров и научных работников. М.: ФизМатЛит, 2006. 816 с.

[16] **Кормен Т., Лейзерсон Ч., Ривест Р.** Алгоритмы: построение и анализ. М.: МЦНМО, 1999. 960 с.

[17] **Корнилов Ф.А., Перевалов Д.С.** Алгоритм поиска структурных изменений на изображениях // СПИСОК-2009: материалы межвузовской научн. конф. по проблемам информатики. Екатеринбург: Изд-во Урал. ун-та, 2009. С. 264–266.

[18] **Корнилов Ф.А., Перевалов Д.С.** Алгоритмы обнаружения структурных изменений на разновременных снимках // Тез. и аннот. докл. региональной научно-технической конф. "Данные дистанционного зондирования Земли для решения практических задач развития территорий". Екатеринбург: ФГУП "Уралгеоинформ", 2009. С. 21–22.

[19] **Корнилов Ф.А., Перевалов Д.С.** Обнаружение изменений объектов земной поверхности по спутниковым снимкам // Проблемы теоретической и прикладной математики. Тез. 41-й всероссийской молодежной конф. Екатеринбург: ИММ УрО РАН, 2010. С. 534–540.

[20] **Корнилов Ф.А., Костоусов В.Б., Перевалов Д.С.** Быстрые алгоритмы анализа изображений со структурными искажениями // Тр. научно-технической конф.-семинара “Техническое зрение в системах управления мобильными объектами-2010”. М.: КДУ, 2011. С. 303–309.

[21] **Корнилов Ф.А., Перевалов Д.С.** Оценка оптимального порога для алгоритма поиска структурных изменений с помощью проектора Пытьева // Современные проблемы математики. Тез. 42-й всероссийской молодежной шк.-конф. Екатеринбург: ИММ УрО РАН, 2011. С. 291–293.

[22] **Корнилов Ф.А., Перевалов Д.С.** Задача обнаружения структурных различий изображений // Алгоритмы и програм. средства парал. вычислений: сб. науч. тр. Екатеринбург: УрО РАН, 2011. Вып. 11. С. 37–56.

[23] **Корнилов Ф.А.** Алгоритм поиска структурных различий на основе Марковских случайных полей // Современные проблемы математики. Тез. международной (43-й всероссийской) молодежной шк.-конф. Екатеринбург: ИММ УрО РАН, 2012. С. 311–313.

[24] **Корнилов Ф.А., Перевалов Д.С.** О применении регуляризации морфологического проектора в задаче поиска структурных различий // Сб. тр. научно-технической конференции “Техническое зрение в системах управления – 2012”. М.: ИКИ РАН, 2012. С. 265–271.

[25] **Корнилов Ф.А.** Исследование алгоритма поиска структурных различий изображений // Тр. Ин-та математики и механики УрО РАН. Екатеринбург: ИММ УрО РАН, 2013. Т. 19, № 1. С. 144–159.

[26] **Корнилов Ф.А.** Теоретическое исследование морфологического проектора Пытьева // Современные проблемы математики. Тез. международной (44-й всероссийской) молодежной шк.-конф. Екатеринбург: ИММ УрО РАН, 2013. С. 322–325.

[27] **Корнилов Ф.А.** Поиск структурных различий изображений: алгоритмы и методика исследования // Машинное обучение и анализ данных, 2014. Т. 1. № 7. С.902–919.

[28] **Корнилов Ф.А.** Поиск структурных различий изображений // Интеллектуализация обработки информации: 10-я международная конференция. Греция, о. Крит, 6–10 октября 2014 г.: Тезисы докладов. — М.: Торус Пресс, 2014. С.118–119.

[29] **Корнилов Ф.А.** Поиск прямоугольных контуров на изображениях с помощью преобразования Хафа // Современные проблемы математики и ее приложений. Труды 45-й

Международной молодежной школы-конференции. Екатеринбург: ИММ УрО РАН, 2014. С. 195–198.

[30] **Корнилов Ф.А.** Исследование влияния рассинхронизации входных изображений на качество работы алгоритмов поиска структурных различий // Машинное обучение и анализ данных, 2015. Т. 1, № 12. С.1686–1695.

[31] **Корнилов Ф.А.** Исследование влияния рассинхронизации входных изображений на качество работы алгоритмов поиска структурных различий // Математические методы распознавания образов: Тезисы докладов 17-й Всероссийской конференции с международным участием, г.Светлогорск, 2015 г. — М.: Торус Пресс, 2015. С.160–161.

[32] **Костоусов В.Б., Кандоба И.Н.** Система автоматизированного дешифрирования космических снимков земной поверхности // Практика приборостроения, 2003. № 1. С. 45–50.

[33] **Костоусов В.Б., Перевалов Д.С.** О задачах обнаружения и локализации объектов на изображении // Тез. докл. междунар. конф. “Алгоритмический анализ неустойчивых задач”. Екатеринбург, 2008. С. 136–137.

[34] Методы компьютерной обработки изображений (под ред. Сойфера В.А.). М.: ФИЗМАТЛИТ, 2003. 784 с.

[35] Пакет для анализа изображений eCognition: `http://www.ecognition.com/`

[36] **Перевалов Д.С.** Исследование алгоритмов обнаружения и локализации объекта на изображении в условиях структурных искажений // Вычислительные технологии, 2009. Т. 14, № 1. С. 94–104.

[37] **Поклад Г.Г.** Геодезия. М.: “Недра”, 1988. 304 с.

[38] Программный продукт для визуализации и обработки данных дистанционного зондирования Земли ENVI: `http://www.exelisvis.com/`

[39] **Пытьев Ю.П.** Морфологический анализ изображений // Докл. АН СССР, 1983. Т. 269, № 5. С. 1061–1064.

[40] **Пытьев Ю.П., Чуличков А.И.** Методы морфологического анализа изображений. М.: ФизМатЛит, 2010. 336 с.

[41] **Сандерс Дж., Кэндрот Э.** Технология CUDA в примерах. Введение в программирование графических процессоров. М.: ДМК Пресс, 2011. 232 с.

[42] **Свешников А.А.** Прикладные методы теории случайных функций. Издание 2-е, переработанное и дополненное. Главн. ред. физ.-матем. лит. из-ва “Наука”, 1968. 464 с.

[43] Сенсор MODIS: `http://modis.gsfc.nasa.gov/about/design.php`

[44] Документация по ГИС: `http://gis-lab.info/docs.html`

[45] Суперкомпьютерный центр “Уран” ИММ УрО РАН: `http://parallel.uran.ru/`

[46] Формат обмена данными MapInfo: `http://mapinfo.forum.hut2.ru/`

[47] **Форсайт Д.А., Понс Ж.** Компьютерное зрение. Современный подход (пер. с англ.). М.: Изд. дом Вильямс, 2004. 928 с.

[48] **Besag J.** On the statistical analysis of dirty pictures // J. Royal Statist. Soc., Ser. B, 1986. Vol. 48. P. 259–302.

[49] **Blaschke T.** Towards a framework for change detection based on image objects // Remote Sensing & GIS for Environmental Studies: Applications in Geography. Gottinger geographische Abhandlungen, 2005. Vol. 113. P. 1-9.

[50] **Borchani M., Clopper F., Atalay V., Stamon G.** Change detection in aerial images // In Proc. 1st Canadian Conf. on Computer and Robot Vision, 2004. P. 1–7.

[51] **Bruzzone L., Prieto D.F.** An Adaptive Semiparametric and Context-Based Approach to Unsupervised Change Detection in Multitemporal Remote-Sensing Images // IEEE Transactions on image processing, April 2002. Vol. 11, № 4. P. 452–466.

[52] **Burnett C., Blaschke T.** A multi-scale segmentation / object relationship modelling methodology for landscape analysis // Ecological Modelling, 2003. 168(3). P. 233–249.

[53] **Burns J.B., Hanson A.R., Riseman E.M.** Extracting straight lines // IEEE Trans. PAMI, July 1986. 8(4). P. 425–455.

[54] **Carlotto M.J.** Detection and Analysis of Change in Remotely Sensed Imagery with Application to Wide Area Surveillance // IEEE Transactions on image processing, January 1997. Vol. 6, № 1. P. 189–202.

[55] Change detection datasets and benchmarks: `http://changedetection.net`

[56] **Chatelain F., Tourneret J-Y., Inglada J.** Change Detection in Multisensor SAR Images Using Bivariate Gamma Distributions // IEEE Transactions on image processing, March 2008. Vol. 17, № 3. P. 249–258.

[57] **Comaniciu D., Meer P.** Mean Shift: A Robust Approach Toward Feature Space Analysis // IEEE Trans. PAMI, May 2002. 24(5). P. 603–619.

[58] **Dai X., Khorram S.** Quantification of the Impact of Misregistration on the Accuracy of Remotely Sensed Change Detection // IEEE Geoscience and Remote Sensing (IGARSS 1997), 1997. Vol. 4. P. 1763–1765.

[59] **Dempster A.P.** A generalization of Bayesian inference // J. Royal Statist. Soc., Ser. B, 1968. Vol. 30, № 2. P. 205–247.

[60] **Elvidge C.D., Yuan D., Ridgeway D.W., Lunetta R.S.** Relative radiometric normalization of Landsat Multispectral Scanner (MSS) data using an automatic scattergram-controlled regression // Photogrammetric Engineering and Remote Sensing, 1995. 61(10). P. 1255–1260.

[61] **Fang C.-Y., Chen S.-W., Fuh C.-S..** Automatic change detection of driving environments in a vision-based driver assistance system // IEEE Trans. Neural Networks, May 2003. Vol. 14, № 3. P. 646–657.

[62] **Fukunaga K., Hayes R.** The reduced Parzen classifier // IEEE Transactions on Pattern Analysis and Machine Intelligence, Apr. 1989. Vol. 11, Issue 4. P. 423–425.

[63] **Gong P.** Change detection using principal component analysis and fuzzy set theory // Canadian Journal of Remote Sensing, 1993. Vol. 19(1). P. 22–29.

[64] **Gray R.M.** Entropy and Information Theory. 2-nd Edition Springer, 2011. 409 p.

[65] **Green D.M., Swets J.M.** Signal detection theory and psychophysics. New York: John Wiley and Sons Ltd., 1966. 455 p.

[66] **Hall O., Hay G.** Multiscale Object-specific Approach to Digital Change Detection // International Journal of Applied Earth Observation and Geoinformation, 2003. 4(4). P. 311–327.

[67] **Hall J., Sturm B.** Radiometric correction of multitemporal Thematic Mapper data for use in agricultural land-cover classification and vegetation monitoring // International Journal of Remote Sensing, 1991. 12(7). P. 1471–1491.

[68] **Haralick R.M.** Statistical and structural approaches to texture // Proceedings of the IEEE, May 1979. Vol. 67, Issue 5. P. 786–804.

[69] **Hay G.J., Marceau D.J., Dube P., Bouchard A.** A multiscale framework for landscape analysis: Object-specific analysis and upscaling // Landscape Ecology, 2001. Vol. 16, Issue 6. P. 471–490.

[70] **Hea C., Zhaoa Y., Weib A.** Land-use/land-cover change detection by using the extended change-vector analysis // 2nd International Conference on Information Science and Engineering (ICISE), 2010. P. 3809–3812.

[71] **le Hegarat-Mascle S., Seltz R., Hubert-Moy L., Corgne S., Stach N.** Performance of change detection using remotely sensed data and evidential fusion: comparison of three cases of application // International Journal of Remote Sensing, 2006. Vol. 27, № 16. P. 3515–3532.

[72] **Hough P.V.C.** Methods, Means for Recognizing Complex Patterns. U.S., Patent 3069654, 1962.

[73] **Hsu Y.Z., Nagel H.H., Rekers G.** New likelihood test methods for change detection in image sequences // Computer Vision, Graphics, and Image Processing, April 1984. Vol. 26, Issue 1. P. 73–106.

[74] **Huang L.K., Wang M.J.** Image thresholding by minimizing the measures of fuzziness // Pattern Recognition, 1995. Vol. 28, № 1. P. 41–51.

[75] **Huang L., Zhang G., Li Y.** An Object-based Change Detection Approach by Integrating Intensity and Texture Differences // IEEE, 2nd International Asia Conference on Informatics in Control, Automation and Robotics. Wuhan, China, 2010. P. 258–261.

[76] **Huertas A., Nevatia R.** Detecting changes in aerial views of manmade structures // Image and Vision Computing, May 2000. Vol. 18, № 8. P. 583–596.

[77] **Kittler J., Illingworth J.** Minimum error thresholding // Pattern Recognition, 1986. Vol. 19, Issue 1. P. 41–47.

[78] **Koeln G., Bissonnette J.** Cross-correlation analysis: mapping landcover change with a historic landcover database and a recent, single-date multispectral image // Proc. 2000 American Society for Photogrammetry & Remote Sensing Annual Convention in Washington, D.C. 9 p.

[79] **Koenderink J. J.** The structure of images // Biological Cybernetics, 1984. Vol. 50(5). P. 363–370.

[80] **Kostousov V.B., Kandoba I.N., Skripnuk V.V, Shabanov G.A.** The system for automated deciphering of cosmic Earth surface photographs // International Archives of Photogrammetry and Remote Sensing, Amsterdam, 2000. Vol. XXXIII, Part B4. P. 425–432.

[81] **Kullback S., Leibler R.A.** On information and sufficiency // The Annals of Mathematical Statistics, 1951. Vol. 22, № 1. P. 79–86.

[82] **Lefebvre A., Corpetti T., Hubert-Moy L.** Object-Oriented Approach and Texture Analysis for Change Detection in Very High Resolution Images // IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2008), July 2008. Vol. 4. P. 663–666.

[83] **Li S.Z.** Markov Random Field Modeling in Image Analysis, 2009. Springer Berlin, Heidelberg. 3rd ed. 362 p.

[84] **Li L., Leung M.K.H.** Robust Change Detection by Fusing Intensity and Texture Differences // Computer Vision and Pattern Recognition. Proc. of the 2001 IEEE Computer Society Conference, 2001. Vol. 1. P. 777–784.

[85] **Li L., Leung M.K.H.** Integrating Intensity and Texture Differences for Robust Change Detection // IEEE Transactions on image processing, February 2002. Vol. 11, № 2. P. 105–112.

[86] **Li W., Li X., Wu Y., Hu Z.** A Novel Framework for Urban Change Detection Using VHR Satellite Images // IEEE The 18th International Conference on Pattern Recognition (ICPR'06). Hong Kong, 2006. P. 312–315.

[87] **Lin C., Huertas A., Nevatia R.** Detecting Buildings from Monocular Images. In Automatic Extraction of Man-Made Objects from Aerial and Space Images // Birkhauser Verlag, Switzerland, 1995. A. Gruen, O. Kuebler and P. Agouris, Editors. P. 125–134.

[88] **Lingg A. J., Zelnio E., Garber F., Rigling B.D.** A Sequential Framework for Image Change Detection // IEEE Transactions on image processing, May 2014. Vol. 23, № 5. P. 2405–2413.

[89] **Liu S.C., Fu C.W., Chang S.** Statistical Change Detection with Moments under Time-Varying Illumination // IEEE Transactions on image processing, September 1998. Vol. 7, № 9. P. 1258–1268.

[90] **Lowe D.G.** Object recognition from local scale-invariant features // Proceedings of the Seventh IEEE International Conference on Computer Vision, 1999. Vol. 2. P. 1150–1157.

[91] **Lu D., Mausel P., Brondizio E., Moran E.** Change detection techniques // International Journal of Remote Sensing, June 2004. Vol. 25, № 12. P. 2365–2401.

[92] **Melgani F., Moser G., Serpico S.** Unsupervised Change Detection Methods for Remote Sensing Images // Optical Eng., 2002. Vol. 41, № 12. P. 81–90.

[93] **Mena J.B.** State of the art on automatic road extraction for GIS update: a novel classification // Pattern Recognition Letters, 2003. Vol. 24, Issue 16. P. 3037–3058.

[94] **Moon T.K.** The expectation-maximization algorithm // Signal Process. Mag., 1996. Vol. 13, № 6. P. 47–60.

[95] **Nielsen A. A., Conradsen K., Simpson J.J.** Multivariate Alteration Detection (MAD) and MAF Postprocessing in Multispectral, Bitemporal Image Data: New Approaches to Change Detection Studies // Remote Sensing of Environment, 1998. Vol. 64. No. 1. P. 1–19.

[96] **Nielsen A. A.** The Regularized Iteratively Reweighted MAD Method for Change Detection in Multi- and Hyperspectral Data // IEEE Transactions on image processing, Vol. 16, No. 2, 2007. P. 463–478.

[97] **Nemmour H., Chibani Y.** Change Detector Combination in Remotely Sensed Images Using Fuzzy Integral // International Journal of Information and Communication Engineering, 2005. 1:4. P. 175–181.

[98] Normalized Difference Vegetation Index: `http://earthobservatory.nasa.gov/Features/MeasuringVegetation/`

[99] **Olsson H.** Regression functions for multitemporal relative calibration of thematic mapper data over Boreal forest // Remote Sensing of Environment, 1993. Vol. 46, Issue 1. P. 89–102.

[100] OriginPro 8.0: `http://www.originlab.com/`

[101] **Otsu N.** A threshold selection method from gray-level histograms // IEEE Trans. Sys., Man., Cyber., 1979. Vol. 9. P. 62–66.

[102] **Perevalov D.S.** Using structural functions for analysis and comparison of algorithms // Proc. of 8-th Open German-Russian Workshop “Pattern recognition and image understanding”. N. Novgorod State Univ., 2011. P. 230–233.

[103] **Pop F., Gruia C., Cristea V.** Distributed algorithm for change detection in satellite images for Grid Environments // IEEE Sixth International Symposium on Parallel and Distributed Computing (ISPDC'07). Hagenberg, Austria, 2007. P. 41–45.

[104] **Radke R.J., Andra S., Al-Kofahi O., Roysam B.** Image change detection algorithms: a systematic survey // IEEE Transactions on image processing, Mar. 2005. Vol. 14, № 3. P. 294–307.

[105] **Robin A., Moisan L., le Hergarat-Mascle S.** An A-Contrario Approach for Subpixel Change Detection in Satellite Imagery // IEEE Transactions on pattern analysis and machine intelligence, November 2010. Vol. 32, № 11. P. 1977–1993.

[106] **Rosin P.** Thresholding for change detection // Proc. of the International Conference on Computer Vision, 1998. P. 274–279.

[107] **Salmon B.P., Kleynhans W., van den Bergh F., Olivier J.C., Grobler T.L., Wessels K.J.** Land Cover Change Detection Using the Internal Covariance Matrix of the Extended Kalman Filter Over Multiple Spectral Bands // IEEE Journal of selected topics in applied earth observations and remote sensing, June 2013. Vol. 6, Issue 3. P. 1079–1085.

[108] **Schott J.R., Salvaggio C., Volchok W.J.** Radiometric scene normalization using pseudoinvariant features // Remote Sensing of Environment, 1988. Vol. 26. P. 1–16.

[109] **Serra J.** Image Analysis and Mathematical Morphology. Academic Press, London, 1982. 610 p.

[110] **Shafer G.** A Mathematical Theory of Evidence. Princeton University Press, 1976. 302 p.

[111] **Skifstad K., Jain R.** Illumination independent change detection from real world image sequences // Computer Vision, Graphics, and Image Processing, 1989. Vol. 46. P. 387–399.

[112] **Song C., Woodcock C.E., Seto K.C., Lenney M.P., Macomber S.A.** Classification and Change Detection Using Landsat TM Data: When and How to Correct Atmospheric Effects? // Remote sens. environ, 2001. Vol. 75. P. 230–244.

[113] **Swaminathan K., Lakshminarayanan G., Seok–Bum Ko.** High Speed Generic Network Interface for Network on Chip Using Ping Pong Buffers // International Symposium on Electronic System Design (ISED), 2012. P. 72–76.

[114] **Tang X.** Texture information in run-length matrices // IEEE Transactions on image processing, 1998. Vol. 7, № 11. P. 1602–1609.

[115] **Tian J., Sun J., Tang Y.** Tricolor Attenuation Model for Shadow Detection // IEEE Transactions on image processing, October 2009. Vol. 18, Issue 10. P. 2355–2363.

[116] **Tomasi C., Manduchi R.** Bilateral Filtering for Gray and Color Images // IEEE Proceedings of the 6-th International Conference on Computer Vision, Bombay, India, 1998. P. 839–846.

[117] **Townshend J.R.G., Justice C.O., Gurney C.** The impact of misregistration on change detection // IEEE Trans. Geosci. Remote Sensing, Sept. 1992. Vol. 30, Issue 5. P. 1054–1060.

[118] **Wang W., Zhao Z.M., Zhu H.** Object-oriented Change Detection Method Based on Multi-scale and Multi-Feature Fusion // Urban Remote Sensing Event, May 2009. P. 1–5.

[119] **Wang L., Chen J., Gong P., Shimazaki H., Tarnura M.** Land cover change detection with a cross-correlogram spectral matching algorithm // International Journal of Remote Sensing, 2009. Vol. 30, Issue 12. P. 3259–3273.

[120] **Wang Z., Bovik A.C., Sheikh H.R., Simoncelli E.P.** Image quality assessment: From error visibility to structural similarity // IEEE Transactions on image processing, Apr. 2004. Vol. 13, Issue 4. P. 600–612.

[121] **Yang X., Lo C.P.** Relative Radiometric Normalization Performance for Change Detection from Multi-Date Satellite Images // Photogrammetric Engineering and Remote Sensing, August 2000. Vol. 66, № 8. P. 967–980.

[122] **Zhao Y., He C.** Improving change vector analysis in Multi-temporal space to detect land cover changes by using cross-correlogram spectral matching algorithm // IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2011. P. 174–177.