

Федеральное государственное автономное образовательное учреждение высшего образования «Уральский федеральный университет имени первого Президента России Б.Н. Ельцина»

На правах рукописи

Дмитриев Никита Владимирович

Повышение точности и производительности алгоритмов анализа и обработки сложноструктурных изображений

2.3.1. Системный анализ, управление и обработка информации, статистика

Автореферат
диссертации на соискание учёной степени
кандидата технических наук

Екатеринбург – 2024

Работа выполнена на кафедре «Мехатроника» ФГБОУ ВО «Уральский государственный университет путей сообщения»

Научный руководитель – кандидат физико-математических наук,
Тарасян Владимир Сергеевич

Официальные
оппоненты:

Вохминцев Александр Владиславович,
доктор технических наук, доцент, ФГБОУ
ВО «Челябинский государственный
университет», заведующий научно-
исследовательской лабораторией
«Интеллектуальные информационные
технологии и системы»;

Митрохин Максим Александрович,
доктор технических наук, доцент, ФГБОУ
ВО «Пензенский государственный
университет», заведующий кафедрой
«Вычислительная техника»;

Ронкин Михаил Владимирович, кандидат
технических наук, ФГАОУ ВО «Уральский
федеральный университет имени первого
Президента России Б.Н. Ельцина», доцент
кафедры информационных технологий и
систем управления Института
радиоэлектроники и информационных
технологий.

Защита диссертации состоится 04 февраля 2025 г. в 11:00 часов на
заседании диссертационного совета УрФУ 2.3.12.13 по адресу:
620062, г. Екатеринбург, ул. Мира, 19, ауд. И-420 (зал Ученого совета).

С диссертацией можно ознакомиться в библиотеке и на сайте ФГАОУ ВО
«Уральский федеральный университет имени первого Президента России
Б.Н. Ельцина»

<https://dissovet2.urfu.ru/mod/data/view.php?d=12&rid=6630>

Автореферат разослан «__» _____ 2024 г.

Учёный секретарь
диссертационного совета



Сафиуллин Николай Тахирович

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования и степень её разработанности.

В настоящее время происходит ускоренное развитие средств обработки и анализа информации вследствие накопления большого числа данных о различных объектах и процессах. Одними из важных составляющих таких средств являются программные комплексы анализа и обработки растровых сложноструктурных изображений, сочетающих в себе числовые и графические данные. Целью таких программных комплексов является получение информации о пространственно-распределённых объектах для дальнейшей обработки. Основными подходами при этом является векторизация и понимание семантических связей между объектами.

Примерами сложноструктурных изображений (ССИ) являются диаграммы, номограммы, чертежи, схемы, архитектурные проекты, карты местностей. В целом данные изображения можно охарактеризовать наличием многих семантических слоёв, наложенных друг на друга в визуальном плане. Например, топографические карты, которые можно назвать типичным представителем ССИ, содержат наиболее полную информацию о местности, поэтому из них можно получить информацию для геоинформационной системы (ГИС), которые можно использовать в качестве поддержки принятия решения на этапе проектирования для решения различных задач, например, оптимального размещения станций мобильной связи, логистических центров, транспортных развязок и дорог различного назначения. Также ГИС являются полезным инструментом анализа для осуществления, например, кадастровых работ, предсказания потенциальных зон пожаров и затоплений. Согласно отчету «*Geographic Information System (GIS) Market*», опубликованному *Allied Market Research*, общемировой рынок ГИС вырастет до 25,5 млрд. долл. к 2030 г., при среднегодовом темпе роста 15,2%.

При этом ручная обработка большого числа ССИ становится экономически неэффективной. Проблема распознавания и анализа образов приобрела выдающееся значение в условиях информационных перегрузок, когда человек не справляется с пониманием поступающих к нему всё возрастающих объёмов данных. В результате его мозг переключается на режим одновременного восприятия и мышления, которому такое распознавание свойственно, но негативно влияет на качество работы в условиях крайней нежелательности совершения ошибок. Поэтому современный период развития информационных технологий характеризуется широким внедрением интеллектуальных технологий в процесс обработки графической информации, а автоматическая обработка и анализ изображений, обладающих сложной структурой, является одним из наиболее перспективных направлений развития.

Решению проблем обработки изображений, анализа и распознавания образов посвящены работы многих исследователей. Основы теории обработки цифровых изображений как частного случая цифровой обработки сигналов, алгоритмы оптической, геометрической и статистической обработки, поиска особенностей изображений были исследованы в работах В.Ф. Кравченко, Н.Н. Красильникова, Ю.П. Пытьева, Я.А. Фурмана, Л.П. Ярославского, J. Canny, R. Gonsales, P. Hough, B. Jähne, D. Marr, T. Pavlidis, W. Pratt, J. Serra, P. Soille, R. Woods.

Вопросам машинного обучения и интеллектуальных методов, в т. ч. искусственных нейронных сетей посвящены труды следующих учёных: В.Н. Вапник, А.И. Галушкин, А.Н. Горбань, В.Д. Мазуров, М.Ю. Хачай, А.Я. Червоненкис, S. Haykin, Y. LeCun, T.M. Mitchel, M. Moller, J. Quinlan, F. Rosenblatt, S. Russe, L. Shapiro, G. Stockman, M. Stone.

Вопросам распознавания образов и компьютерного зрения внимание уделено в работах Ю.В. Визильтера, А.Л. Горелика, И.Б. Гуревича, С.Ю. Желтова, Ю.И. Журавлёва, В.В. Сергеева, В.А. Сойфера, С.М. Bishop, R. Duda, D. Forsyth, P. Hurt, E. Patrick, J. Tou, J. Ward.

Применительно к обработкам топографических карт и других ССИ можно выделить исследования С.А. Барталева, Д.Ю. Васина, Ю.Г. Васина, А.В. Дунаевой, В.Т. Калайды, Ю.Н. Кунакова, С.Ю. Мирошниченко, В.П. Пяткина, О.А. Славина, А.В. Никонорова, М.В. Ронкина, В.С. Титова, Р.А. Томаковой, Л.И. Черновой, R. Szendrey, H. Chen, J. Pouderoux, S. Leyk, Y.-Y. Chiang, T. Liu.

Существующие программные решения, которые помогают автоматизировать часть операций по обработке цифровых изображений, реализующих научные результаты вышеуказанных исследователей и собственные коммерческие разработки: ГИС «Панорама», *ArcGIS*, *ArcMap*, *Easy Trace*, *MapEdit*, *MapInfo*, *QGIS*, *Spot-Light* и другие.

Недостатками существующих программных решений и научных работ является недостаточное внимание к вопросам точности, производительности и условий применимости создаваемых алгоритмов и реализуемых интеллектуальных моделей. И если ранее данная проблема не стояла так остро из-за слаборазвитого рынка компаний, занимающихся цифровой обработкой ССИ, то сейчас это приводит к необходимости закупки дорогого аппаратного обеспечения и формирования штата сотрудников, выполняющих монотонные операции по поиску и распознаванию образов с малой степенью автоматизации. Без дальнейшей научной проработки экономический рост может быть затруднён, что свидетельствует об актуальности выбранной темы исследования.

Целью исследования является повышение точности и производительности алгоритмов анализа и обработки сложноструктурных изображений с помощью объединения различных подходов обработки изображений.

Задачи, решение которых необходимо для достижения цели исследования:

1. Анализ современного состояния и развития подходов обработки ССИ с целью выработки методики построения программных систем анализа и обработки ССИ, с помощью которой можно повысить точность и производительность.

2. Разработка обобщённой методики анализа и обработки ССИ и разработка гибридных алгоритмов сегментации, локализации, распознавания и группировки образов различной размерности на ССИ, используя классические и интеллектуальные модели и методы, опираясь на критерии качества и быстродействия.

3. Экспериментальная проверка разработанных алгоритмов на различных классах ССИ с вычислением критериев качества, быстродействия и стабильности, сравнение полученных результатов с современными научными исследованиями.

4. Разработка специализированного алгоритмического и программного обеспечения для анализа и обработки цифровых топографических карт, как типичного представителя ССИ, содержащего все виды объектов по размерности: точечные, линейные, площадные.

Объектом исследования являются методы и алгоритмы анализа и обработки растровых сложноструктурных изображений.

Предметом исследования являются показатели качества алгоритмов анализа и обработки сложноструктурных изображений, такие как точность и производительность.

Методология и методы исследования. Для решения поставленных в диссертационной работе задач были применены методы интеллектуальной обработки данных, в т. ч. обработки данных с использованием искусственных нейронных сетей, методы обработки изображений, в т. ч. методы математической морфологии, методы геометрической и семантической фильтрации, числовые методы линейной алгебры, методы оптимизации на основе генетических алгоритмов.

Научная новизна диссертационной работы определяется следующими новыми научными результатами:

1. На основе проведённого анализа современного состояния и развития подходов обработки цифровых изображений, предложена классификация сложноструктурных изображений, отличающихся использованием явных и неявных правил отображения распределённых образов различной размерности, и методика построения алгоритмического и программного обеспечения анализа и обработки таких изображений (соответствует п. 7 паспорта специальности).

2. Разработаны алгоритмы сегментации, локализации, распознавания и группировки точечных, линейных и площадных объектов на сложноструктурных изображениях, основанные на комплексном

использовании известных методов анализа и обработки изображений (соответствует пп. 4, 5 и 12 паспорта специальности).

3. Разработан алгоритм комплексного анализа распознанных числовых отметок высот и горизонталей цифровых топографических карт, отличающийся представлением контурной карты в виде ортура и способный устранять ошибки предшествующего этапа распознавания (соответствует пп. 4, 5, 12 паспорта специальности).

Теоретическая значимость диссертационного исследования заключается в разработке классификации ССИ и методики построения алгоритмического и программного обеспечения их анализа и обработки с помощью гибридного, итеративного и комплексного подхода.

Практическая значимость работы заключается в экспериментальной проверке разработанной методики обработки ССИ и набора алгоритмов в виде программных модулей. При этом выполняются следующие пункты.

1. Предложенная методика построения программных систем анализа и обработки сложноструктурных изображений позволяет производить обработку изображений с качеством 200 dpi с высокой точностью и производительностью, что приводит к снижению требований на аппаратное обеспечение персональных компьютеров и сокращению времени работы пользователей.

2. Разработанные алгоритмы сегментации, локализации, распознавания и группировки точечных объектов ССИ позволяют понизить ошибку в распознавании цифр относительно существующих методов и понизить количество вычислений за счёт сокращения рассматриваемых пикселей.

3. Разработанные алгоритмы локализации, распознавания и группировки линейных объектов ССИ позволяют понизить количество разрывов и повысить производительность вычислений относительно существующих методов.

4. Разработанные алгоритмы сегментации и распознавания площадных объектов ССИ позволяют использовать более простые интеллектуальные модели с последующей обработкой относительно сложных, что понижает требования к аппаратному обеспечению.

5. Предложенный алгоритм комплексного анализа распознанных точечных и линейных объектов цифровой топографической карты позволяет получить пропущенные данные и исправить неправильно распознанные отметки относительно других существующих методов и программных решений, в которых данные операции выполняются пользователем.

Научные и практические результаты диссертационной работы были реализованы при выполнении научного проекта РФФИ № 17-08-01123 А «Математическое моделирование транспортных сетей в условиях развития

высокоскоростного наземного транспорта» и в рамках госбюджетной научно-исследовательской работы «МХТ-8» по теме «Интеллектуальное управление и анализ данных в мехатронных и робототехнических системах».

Положения, выносимые на защиту:

1. Предложенная методика построения программных систем анализа и обработки ССИ формализует данный процесс благодаря разделению обрабатываемых объектов по размерностям, а процесса – по этапам, включая этапы внутри- и межсемантического анализа, что позволяет унифицировать проектирование алгоритмического и программного обеспечения для различных классов ССИ, учитывая их особенности.

2. Набор разработанных алгоритмов, в т. ч. этапов сегментации, локализации, распознавания и группировки объектов ССИ различной размерности, позволяет повысить быстродействие и точность указанных этапов обработки за счёт оптимизации получения параметров алгоритмов, последовательностей операций и гиперпараметров интеллектуальных моделей, относительно существующих подходов, в которых данные параметры и модели подбираются вручную.

3. Разработанное программное обеспечение для анализа и обработки ССИ подтверждает применимость предложенной методики и набора алгоритмов на примере смоделированных навигационных морских карт, спутниковых снимков и сканированных изображений топографических карт в разрешении 200 dpi, что на более детальном разборе последних позволяет сократить время автоматической обработки одной карты с целью получения ЦМР до ~10 минут, необходимой ручной обработки до 20..60 минут с полностью верным восстановлением значений высот горизонталей, по сравнению с использованием существующих программных средств, где обработка одной топографической карты может занимать до одного рабочего дня с возможностью ошибки, вы-званной ручным трудом.

Степень достоверности полученных в диссертационной работе результатов и выводов обеспечивается логической обоснованностью используемых подходов к анализу и обработке изображений, их грамотном сочетании, обеспечивающем повышение точности и производительности; подтверждается воспроизводимостью и непротиворечивостью данных, полученных в ходе вычислительных экспериментов, получением свидетельств о государственной регистрации программ для ЭВМ, внедрением полученных результатов, апробацией на международных конференциях, наличием публикаций основных результатов работ в рецензируемых российских и зарубежных изданиях.

Внедрение результатов диссертационного исследования. Результаты диссертационного исследования прошли апробацию и были внедрены в деятельности АО «Урало-Сибирской Гео-Информационной

компаний». Результаты исследования внедрены в учебный процесс и научно-исследовательскую работу Уральского государственного университета путей сообщения при проведении учебных занятий с бакалаврами по направлениям «Мехатроника и робототехника», «Информатика и вычислительная техника». Соответствующие акты прилагаются к диссертации.

Апробация работы. Основные результаты диссертации докладывались и обсуждались на следующих конференциях и семинарах: «Математические методы решения исследовательских задач» (г. Екатеринбург, 2014 г., 2015 г.), «Информационные технологии и нанотехнологии» (г. Самара, СГАУ, 2016 г.), «Математическое моделирование и информационные технологии» (г. Екатеринбург, 2016 г.), «Управление и обработка информации в сложных системах» (г. Екатеринбург, 2017 г.), выставка «ИННОПРОМ» (г. Екатеринбург, 2017 г.), «Applied System Analysis and Computer Science» (г. Екатеринбург, УрФУ, 2018 г.), «Научный семинар по информационным технологиям» (г. Челябинск, НИУ ЮУрГУ, 2018) «Современные проблемы математики и ее приложений» (г. Екатеринбург, ИММ УрО РАН, 2019 г.), «Applications of Mathematics in Engineering and Economics» (г. Созополь, Болгария, 2019 г.), «Инновационные производственные технологии и ресурсосберегающая энергетика» (г. Омск, ОмГУПС, 2021), «TLC2M Transport: Logistics, Construction, Maintenance, Management» (г. Екатеринбург, 2022).

Публикации. По теме диссертации опубликовано 17 работ, из них 5 статей опубликованы в рецензируемых научных изданиях, определенных ВАК РФ и Аттестационным советом УрФУ, включая 1 статью в издании, индексируемом в международной цитатно-аналитической базе Scopus; имеется 2 свидетельства о государственной регистрации программы для ЭВМ.

Личный вклад. Автором проанализирована и разработана концепция преобразования изображения топографических карт в объёмную модель местности, спроектирован прототип программного комплекса, произведена разработка и экспериментальная проверка алгоритмов сегментации, поиска, распознавания и группировки объектов на изображениях, алгоритма восстановления плана горизонталей и интерполяции. Опубликованные работы подготовлены диссертантом самостоятельно, все основные положения тезисов докладов конференций также сформулированы лично автором, научному руководителю принадлежит постановка задачи исследования.

Структура и объём диссертации. Диссертация состоит из введения, четырёх глав, заключения, библиографии и трёх приложений. Общий объём диссертации 222 страницы, включая 35 рисунков и 40 таблиц. Библиография включает 273 наименования.

СОДЕРЖАНИЕ РАБОТЫ

Во **введении** обосновывается актуальность темы диссертации, формулируется цель и задачи исследования, обсуждается новизна и практическая ценность выносимых на защиту результатов, даётся краткая характеристика содержания работы.

В **первой главе** производится анализ современного состояния проблемы анализа и обработки ССИ. Так как термин не обладает к настоящему моменту строгого определения, поэтому были рассмотрены свойства изображений, относимых современными исследователями к обладающими сложную структуру. К ним были отнесены: разнообразие образов, сочетание графических (в т. ч. различной размерности) и текстовых элементов, большеформатность, зависимость атрибутов объектов от их местоположения, гетерархичность, нарушение и/или нерегулярность правил отображения образов, высокая информационная ёмкость, наличие неоднозначной, неполной, косвенной, вероятностной информации. Было выявлено наличие у ССИ визуального плана, состоящего из пикселей растра, и семантического плана, содержащего информацию об объектах, расположенных на ССИ.

В итоге ССИ было определено следующим образом:

Сложноструктурное изображение – растровое изображение, которое содержит объекты, относящиеся к различным семантическим планам, и которое обладает набором явных и неявных правил отображения этих объектов на визуальный план с помощью различных техник в виде распределённых образов, подразумевающих возможность однозначного восстановления смысла исходных объектов и их атрибутов.

Также разработана классификация ССИ, опирающаяся на природу их получения: непосредственно отражающие реальность (I класс), созданные человеком рукотворно (II класс), созданные человеком автоматизировано (III класс). Классы представлены вместе с десятью подклассами и современными научными исследованиями, методами и алгоритмами, применяющимися для их обработки.

Далее проведён анализ теории обработки изображений, её концепции и типичные задачи, оптические, геометрические и интеллектуальные подходы к их решению. Различные подходы сравниваются и делается предположение о необходимости разработки гибридных методов, сочетающих в себе быстрдействие оптических методов и качество работы интеллектуальных.

Процесс обработки ССИ разобран по всем основным этапам: предобработке, локализации и распознавании образов, анализа сцены и постобработке. Выделены основные критерии качества, с помощью которых можно оценивать алгоритмы анализа и обработки ССИ,

основанные на ошибках первого и второго рода, мере Жаккара, стабильности машинного обучения интеллектуальных моделей.

После анализа современного состояния проблемы анализа и обработки ССИ были сделаны следующие выводы.

1. Использование конкретных методов может давать хороший результат при обработке изображений, имеющих узкую направленность. При этом ССИ тяготеют к консолидации графики различного типа и гетерархичности семантических планов, что приводит к тому, что необходимо использовать гибридные и итеративные алгоритмы.

2. Большинство исследований рассматривает только отдельные шаги обработки ССИ, без комплексного решения задачи, оставляя эту работу оператору. Без решения комплексных задач невозможна дальнейшая автоматизация и автоматический поиск ошибок. Анализ неявных правил отображения объектов или анализ влияния нарушения этих правил выражен лишь в малой части работ.

3. В большинстве исследований не отражены временная сложность алгоритмов, конкретное время выполнения и аппаратное обеспечение. При известности времени оно обычно является достаточно большим из-за использования сложных интеллектуальных моделей и неоптимальных алгоритмов полиномиального времени высокого порядка. Точность и стабильность алгоритмов также являются важными критериями качества, но именно суммарное количество времени работы ЭВМ и оператора является ключевым в плане эффективности.

4. Многие исследования тяготеют к рассмотрению более простых моделей ССИ, которые удобно обрабатываются отдельными, негибридными методами. Реальные ССИ часто могут кроме своей сложности иметь низкое разрешение, высокую зашумлённость, неоднородность фона и факты нарушения правил отображения. Всё это приводит к высокой степени дробления сегментов, что говорит о необходимости создания алгоритмов, которые могут восстанавливать целостность образов через группировку.

5. Качество разработанных интеллектуальных моделей показано в виде конечных значений, что не позволяет оценить их стабильность при многократном запуске из-за стохастичности алгоритмов машинного обучения. Стабильными по определению являются линейный регрессор, k - MN классификатор и SVM , для всех типов ANN анализ стабильности является открытой проблемой.

6. Часто в исследованиях используется простейшая пороговая обработка, многие важные параметры методов и моделей задаются константами, выбор значений размеров скользящих окон и структурных элементов не обоснован. Всё это приводит к слабой гибкости разработанных алгоритмов. Данный подход целесообразен при

использовании интеллектуальных моделей с большим количеством решающих элементов, которые поэтому и могут быть простыми.

7. Материалы ССИ, использующиеся для обучения и тестирования интеллектуальных моделей и алгоритмов, либо имеют слабую разметку данных, либо представляют собой только куски ССИ. Почти во всех работах тренировочный набор данных скорее аугментируется с использованием стандартных операций (симметрия, вращение, сдвиг, добавление шума), что приводит к необходимости рассмотрения более сложных интеллектуальных моделей, вместо попытки упрощения данных и нахождения их инвариантов, в принципе невосприимчивых к подобным операциям.

Во **второй главе** предложена обобщенная методика анализа и обработки ССИ (рис. 1), основанная на последовательном преобразовании растрового изображения в набор точечных (ТО), линейных (ЛО) и площадных (ПЛО) объектов с их описаниями, пригодными для дальнейшего использования формальными методами.

В соответствии с методикой были разработаны отдельные алгоритмы, которые могут быть использованы для обработки ССИ любого типа с любой интеллектуальной моделью в качестве решателя.

Алгоритм нахождения информации о соседних пикселях является альтернативой широко применяемого метода скользящего окна, когда в качестве центра окон рассматриваются все возможные пиксели растра; в отличие от него разработанный алгоритм принимает на вход информацию о необходимых центрах окон *Points* и угол, под каким окно $a_x \times a_y$ должно быть расположено. Временная сложность данной функции соответствует $\Theta(|Points| * a_x * a_y)$.

Алгоритм многометочной сегментации использует паноптический подход к сегментации, при котором при помощи простейших вычислителей над скользящими окнами пространство растра разбивается на сегменты-суперпиксели. Отличием алгоритма является сохранение информации о мере принадлежности пикселя рассматриваемым классам K вместо строгой классификации. При использовании *ANN* прямого распространения с одним скрытым слоем, состоящим из N нейронов, временная сложность алгоритма для растра площади S будет равна $\Theta(S * N * (a^2 + K))$.

Алгоритм локализации ТО и ЛО представляет собой набор операций математической морфологии, пороговой обработки и фильтрующих операций по свойствам объектов.

Генетический алгоритм получения последовательности операций для локализации основывается на представлении такой последовательности в виде хромосомы (набора генов), которые модифицируются с помощью генетических операторов скрещивания и мутации.

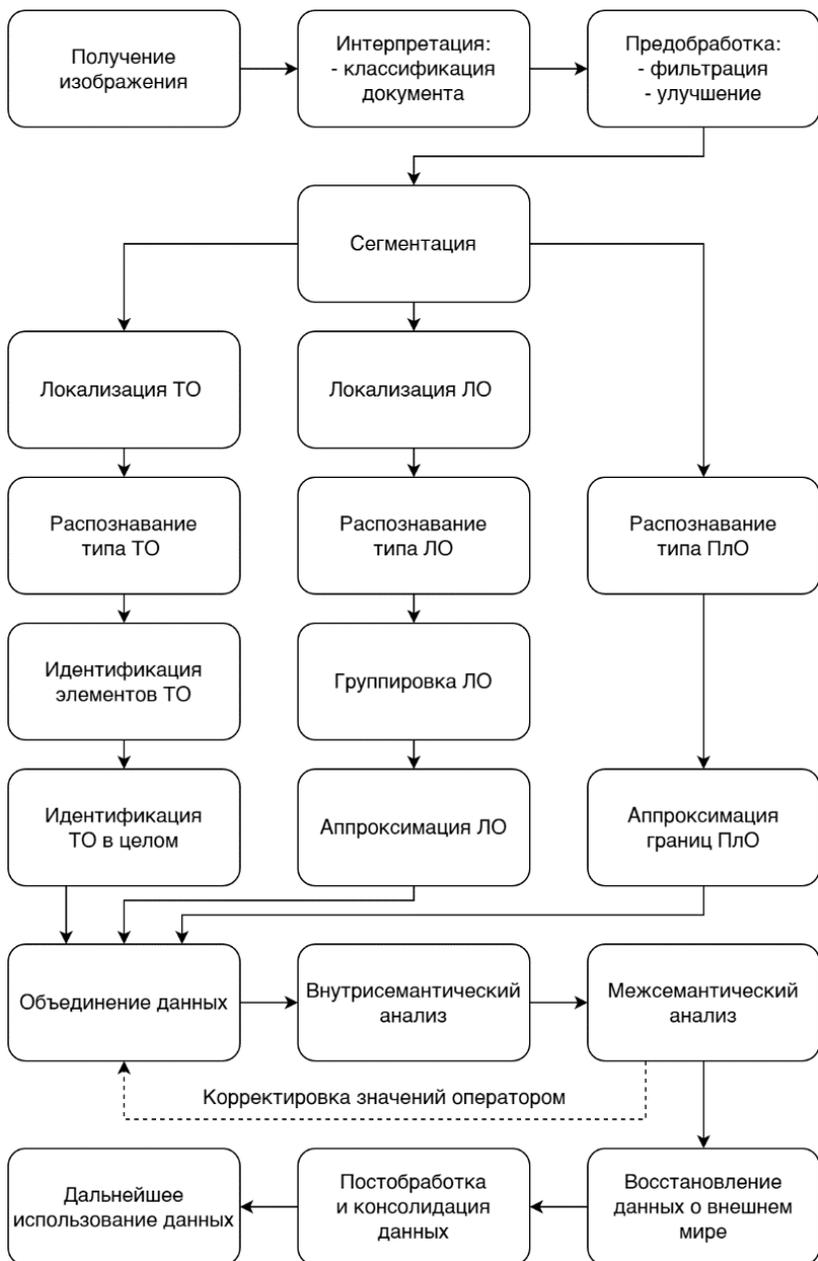


Рис. 1 – Обобщённая методика программного анализа и обработки ССИ

В качестве функции приспособленности используются формулы:

$$RecallGA(P') = [|P' \wedge P_{crit}| \geq |P'_\pi \wedge P_{crit}| \text{ or } |P' \wedge P_{crit}| \geq 0.95 * |P_{crit}|], \quad (1)$$

$$FitnessFunction(P') = RecallGA(P') * \frac{|P' \wedge P_{crit}|}{|P_{crit}|} * \frac{|P' \wedge P_{ideal}|}{|P' \vee P_{ideal}|}, \quad (2)$$

где P' – бинарное изображение, полученное в результате выполнения операций над оригинальным в текущей хромосоме, P_{crit} – бинарное изображение, где единицами отмечены центры размеченных фигур, P'_π – бинарное изображение, полученное для предка данной хромосомы, $[statement]$ – скобка Айверсона, $|P|$ – мощность ненулевых пикселей бинарного изображения, P_{ideal} – размеченное бинарное изображение.

Под скрещиванием понимается такое определение потомка-хромосомы, при котором от одного предка берётся левая часть (до пороговой операции), а от другого – правая (после пороговой).

В качестве мутации рассматривалось 3 варианта: вставка гена-операции в хромосому, модификация случайного гена, удаление случайного гена (кроме пороговой обработки). Таким образом, построенный генетический алгоритм отличается от существующих возможностью изменения числа генов в хромосоме.

Алгоритм классификации ПЛО основан на принципе локальности, который можно выразить в существовании такой функции f_{local} , определяющей класс пикселя по классам пикселей его соседей и удовлетворяющей следующему условию для почти каждого пикселя:

$$class(p) = f_{local}(\{class(p') \mid d(p, p') \leq d_{max}\}), \quad (3)$$

где $class(p)$ – класс фонового изображения текущего пикселя p , определяемого через свои координаты раstra, p' – пиксели, для которых выполняется условие соседства, d – метрика, по которой считается расстояние между двумя пикселями, d_{max} – максимальное значение данной метрики.

Алгоритм отличается использованием информации о соседних пикселях в окне $a \times a$ в виде гистограмм с количеством столбцов $bins$: причём сначала используются гистограммы по цветовым координатам, а на последующих итерациях – по найденным классам K . Временная сложность зависит от выбранной интеллектуальной модели, может быть оценена как $O(S * (G * N_{recon} + a^2 + bins * K) + S' * (a^2 + K) * N_{iter})$, где G – размер фильтра Гаусса, N_{recon} – количество итераций размывтия в функции восстановления, S' – количество неклассифицированных пикселей после первого этапа алгоритма, N_{iter} – число итераций второго этапа, $N_{iter} = O(\sqrt{S'})$.

Двухкаскадный алгоритм распознавания ТО основывается на объединении классического подхода распознавания графических образов,

выраженного в виде анализа пикселей из скользящего окна, и подхода, опирающегося на анализ признаков.

Распознаваемые образы разбиваются на классы и подклассы, для класса в целом и для подклассов обучаются интеллектуальные модели. Аналогично алгоритму сегментации с выхода первой модели, использующейся для распознавания подклассов образов, берётся мера принадлежности к подклассу для точек, определённых на этапе локализации. Все меры перемножаются на выход второй модели, отвечающей за распознавание классов и после этого происходит вычисление подкласса с помощью функции $argmax$.

Выход первого каскада представляет собой набор *суперсегментов*. Суперсегмент определяется нами как непрерывное множество пикселей на индексированном изображении, где значение цвета пикселя определяется индексом класса. Суперсегменты проходят через операции фильтрации, необходимые для их отделения и отбрасывания ложных суперсегментов. Второй каскад алгоритма получает на вход интеллектуальной модели доли исходных подклассов класса внутри суперсегмента в качестве набора признаков. Таким образом, второй каскад алгоритма является более абстрактным, что позволяет использовать его отдельно от первого каскада в случае, когда образы могут менять размеры из-за масштабирования.

Временная сложность алгоритма без учёта выбора интеллектуальной модели соответствует $\Theta(N_{points} * a_x * a_y + S * K)$.

Алгоритм группировки ТО представляет собой кластеризацию ТО с использованием геометрического хэширования и выполнения последующих фильтрующих операций. В качестве фильтрации используется геометрический анализ, основанный на сравнении геометрических параметров с образцовыми, и семантический анализ, использующий информацию о распознанных ранее образах.

Алгоритм группировки ЛО основан на выполнении последовательности соединения n концевых точек скелетизированного изображения, полученного на выходе этапа сегментации для ЛО. Для проверки возможности соединения были выделены 5 условий: локальности, согласования углов касательных, согласования касательной с углом наклона линии соединения (проверяются по всем возможным парам, поэтому для условий I-III время работы $\Theta(n^2)$), доминирования, максимального правдоподобия (проводятся после кластеризации концевых точек, время работы $\Theta(N_{pairs})$, $N_{pairs} \ll n^2$).

Под касательной в точке на ЛО растра подразумевается значение, полученное по окрестности данной концевой точки:

$$\alpha_i = -atan2 \left(\frac{\sum_{(x,y) \in S_i} y}{\sum_{(x,y) \in S_i} 1} - y_i, \frac{\sum_{(x,y) \in S_i} x}{\sum_{(x,y) \in S_i} 1} - x_i \right), \quad (4)$$

где i – индекс точки, S_i – множество пикселей концевой участка линейного сегмента около данной точки, $atan2(y, x)$ – функция четырёхквadrантного арктангенса, получающего на вход координаты вектора, исходящего из центра координат.

Для условия максимального правдоподобия используется мера:

$$measure(e_i, e_j) = (|\alpha_i - \gamma_i| + |\alpha_j - \gamma_j|) * D(e_i, e_j), \quad (5)$$

где α_i, α_j – касательные в точках i и j , γ_i – угол наклона линии соединения, $D(e_i, e_j)$ – расстояние между концевыми точками i и j .

В **третьей главе** производится экспериментальная проверка разработанных алгоритмов для двух подклассов ССИ: Пб2 и Ia.

В качестве примера класса Пб2 для тестирования алгоритмов, связанных с распознаванием ТО, использовались смоделированные морские навигационные карты. Моделирование основывалось на создании псевдопланетарного рельефа с помощью квазифрактального алгоритма *diamond-square* и затем создании на его основе 20 моделей морских навигационных карт на основе неполного следования ГОСТ Р 58251-2018. Для цифр отметок глубин использовался датасет *MNIST*, цифры располагались на разном расстоянии, в т. ч. с пересечением. Были добавлены другие точечные, линейные и площадные объекты; проведено размытие и зашумление изображения, имитация сканирования с разрешением 200 dpi.

На рис. 2 показана общая схема обработки, основанная на разработанной методике, на рис. 3 – процесс преобразования участка карты. Итоговая точность распознавания отметок глубин 89,1%, время работы: 18 сек. CPU на одну карту размера 2044×2044 пикселей.

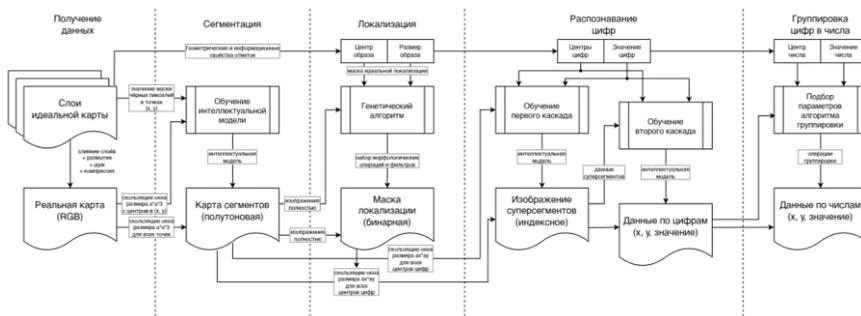


Рис. 2 – Схема обучения и тестирования обработки точечных объектов

В качестве изображений класса Ia для тестирования алгоритмов распознавания ПЛО, использовались аэрофотоснимки из соревнования *DeepGlobe 2018 Land Cover Classification*. В датасете каждому из 803 изображений размера 2448×2448 была поставлена в соответствие маска,

пиксели которой принадлежали к одному из семи классов: *urban*, *agriculture*, *rangeland*, *forest*, *water*, *barren*, *unknown*. В качестве метрики качества использовалось среднеклассовое значение *mIoU*:

$$mIoU = \frac{1}{k} * \sum_{j=1}^k IoU_j = \frac{1}{k} * \sum_{j=1}^k \frac{\sum_{i=1}^n TP_{ij}}{\sum_{i=1}^n TP_{ij} + \sum_{i=1}^n FP_{ij} + \sum_{i=1}^n FN_{ij}}, \quad (6)$$

где IoU_j – значение метрики IoU для j -го класса по точкам всех n изображений, TP_{ij} – количество верно классифицированных точек к классу j на изображении i , FP_{ij} – количество неверно классифицированных точек к классу j , FN_{ij} – количество неверно классифицированных точек к любому другому классу, кроме j .

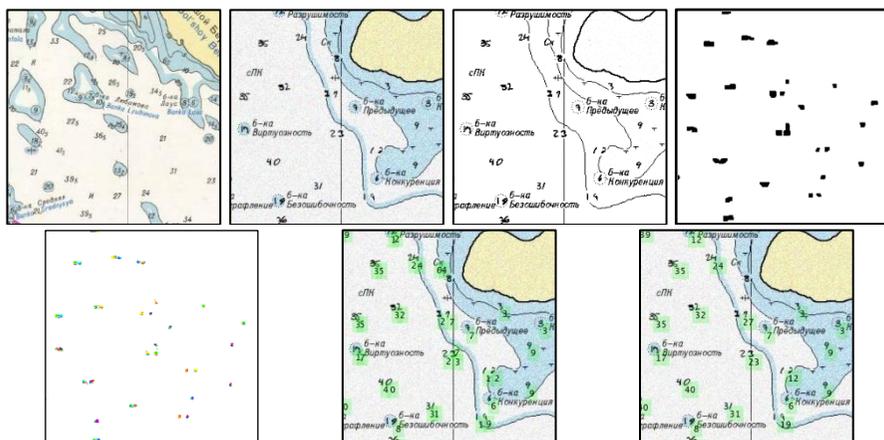


Рис. 3 – Пример преобразования участка смоделированной навигационной карты, слева-направо, сверху-вниз: реальная карта, смоделированная карта, выход алгоритма сегментации, выход алгоритма локализации, выход первого каскада алгоритма распознавания, выход второго каскада алгоритма распознавания, выход алгоритма группировки

В качестве интеллектуальной модели использовались *ANN*, выходные меры классов которых корректировалась с помощью поправочных коэффициентов для каждого класса. Дальнейшая постобработка была связана с удалением небольших фрагментов и последовательным выполнением операции моды по окнам различных размеров. Значения параметров подбирались с помощью МНК и подбором из степенных рядов. Анализ результатов производился с помощью *PCA*.

Результаты эксперимента показали улучшение точности работы алгоритмов даже при использовании *ANN* прямого распространения относительно такой *DNN* как *DeepLabV3+* (таблица 1, звездочкой

отмечены подходы с постобработкой, жирным предложены алгоритмы), при этом результат для *DNN* также удалось улучшить с помощью примененной методики. Результаты экспериментов изображены на рис. 4.

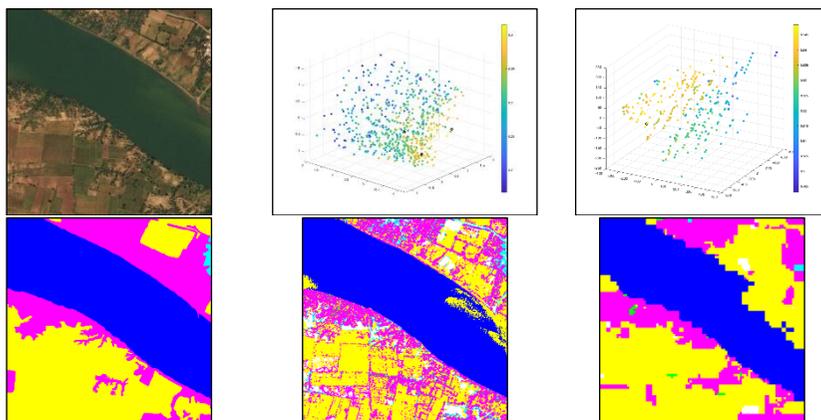


Рис. 4 – Пример преобразования аэрофотоснимка, слева-направо, сверху-вниз: оригинальный снимок, *PCA*-анализ подбора коэффициентов, *PCA*-анализ подбора параметров фильтрации, исходные размеченные данные, выход ИНС, выход с фильтрацией

Таблица 1 – Показатели работы алгоритмов для обработки датасета *DeepGlobe 2018 Land Cover Classification*

Подход	<i>mIoU</i> , %
<i>Demir</i>	43,3
<i>FCN-32s</i>	45,9
<i>DeepLabv3</i>	46,8
<i>DeepLabv3*</i>	51,0...52,7
<i>ANN*</i>	56,1
<i>DeepLabv3+</i>	56,2
<i>DeepLabv3+*</i>	62,0

Четвёртая глава посвящена разработке и тестированию ПО автоматизированной обработки цифровых топографических карт (ЦТК), которые можно отнести к подклассу IIIб. В качестве основной решалась задача получения цифровой модели местности (ЦМР) из растровых ЦТК, поэтому использовались разработанные алгоритмы для обработки ТО и ЛО, т. к. основными элементами, отражающими рельеф, являются отметки высот, горизонталы, полугоризонталы, подписи горизонталей. Дополнительно

было разработано ещё два, отражающих специфику данных ССИ.

Алгоритм комплексного анализа данных ЦТК заключается в совместной обработке данных, относящихся к рельефу, его особенностью является представление плана горизонталей в виде орграфа, в котором вершинами являются значения высот горизонталей, а рёбрами – возможность перехода с одной горизонтали на другую без пересечения других горизонталей (отношение соседства). Отношение соседства и

высоту текущей горизонтали можно вычислить, используя морфологические операции над полупространствами, образованными линиями текущей и соседней горизонталей:

$$z_j = \begin{cases} z_i + \left((N_{in,i} = N_{in,j}) - (N_{out,i} = N_{out,j}) \right) * s_i * step, s_i = -s_j \\ z_i + \left((N_{in,i} = N_{out,j}) - (N_{out,i} = N_{in,j}) \right) * s_i * step, s_i = +s_j \end{cases}, \quad (7)$$

где z_i – относительная высота i -ной горизонтали, $N_{in/out,i}$ – номер соседнего подпространства, находящегося внутри или снаружи i -ной горизонтали, s_i – значение ската i -ной горизонтали, $step$ – значение шага горизонталей.

При известности плана горизонталей и направления их скатов, отображаемых в виде бергштрихов, можно восстановить все значения высот горизонталей, зная только одно значение, что можно реализовать, выполняя обход графа в ширину с выполнением расчёта значения высоты относительно корневой вершины.

Так как при распознавании подписей горизонталей могут быть допущены ошибки, то восстановить из любого значения план нельзя, поэтому сначала нужно произвести вычисление моды значения корневой горизонтали по существующим данным, что при вероятности ошибки меньшей 50% может дать правильный результат. Таким образом алгоритм не только способен восстановить план горизонталей, но и избавиться от ошибки неправильного распознавания подписей и отметок высот.

Алгоритм интерполяции, учитывающий специфику ЦТК, должен лучше передавать рельефность. Для него все точки карты разбивались на три типа: находящиеся между двух горизонталей различной высоты, расположенными внутри замкнутой горизонтали локального экстремума высоты или между границей карты и ближайшей горизонталью. Для точек первого типа была разработана следующая формула, учитывающая отдельный расчёт весовых коэффициентов для ближайшей и дальней горизонталей:

$$Z_0 \approx \frac{Z_1 * \sum_{i=1}^n \frac{1}{d_{i,0}} + \frac{Z_2 * n}{D_{2,0}}}{\sum_{i=1}^n \frac{1}{d_{i,0}} + \frac{n}{D_{2,0}}} = \frac{Z_1 * \sum_{i=1}^n \frac{D_{2,0}}{n * d_{i,0}} + Z_2}{\sum_{i=1}^n \frac{D_{2,0}}{n * d_{i,0}} + 1} = \frac{\rho * Z_1 + Z_2}{\rho + 1}, \quad (8)$$

где Z_1, Z_2 – высоты первой и второй горизонталей, $d_{i,0}$ – расстояние от i -ной точки ближайшей горизонтали до текущей, $D_{2,0}$ – среднее расстояние от текущей точки до дальней горизонтали, $\rho = \sum_{i=1}^n \frac{D_{2,0}}{n * d_{i,0}}$ – среднее отношение расстояния ко второй горизонтали к расстоянию до точек первой.

Точки второго и третьего типа рассчитываются с учётом информации о градиенте, полученной от точек первого типа:

$$\alpha(r_i) = \alpha_0 - \mu * r_i, \quad (9)$$

где r_i – расстояние от точки на горизонтали в сторону точки экстремума или границы карты, α_0 – значение крутизны склона в точке на горизонтали, μ – линейный коэффициент.

После реализации всех алгоритмов в программной среде *Matlab* был получен программный комплекс (ПК) для преобразования ЦТК в ЦМР, состоящий из семи программных блоков, каждый из которых отвечает за свой этап преобразования.

Для получения параметров алгоритмов, обучения интеллектуальных моделей и тестирования ПК использовались полноцветные топографические карты 1965-1982 года масштабов 1 : 50 000 и 1 : 25 000 с разрешением 200-300 dpi, находящимся в перечне открытых данных.

В процессе обучения исследовались следующие интеллектуальные модели: ANN прямого распространения (как распознающие все классы, так и отдельные), метод k ближайших соседей, метод опорных векторов, метод решающих деревьев.

Сегментирование производилось для «черных» (I) и «коричневых» (II) объектов, при этом лучшей с точки зрения качества (F_I -мера = $98,3 \pm 0,2$, F_{II} -мера = $96,3 \pm 0,2$), быстродействия ($T = 6,6$ с/Мп) и стабильности была выбрана ANN прямого распространения со скользящим размером окна равном пяти и 81 нейроном на скрытом слое.

С помощью генетического алгоритма была получена оптимальная последовательность морфологических и фильтрующих операций для локализации центров цифр отметок высот и подписей горизонталей. При потере 0,1% цифр класса I и 0% цифр класса II количество локализованных пикселей было сокращено в 2,6 и 6,1 раз соответственно относительно операций, подобранных вручную.

Двухкаскадный алгоритм распознавания (ANN прямого распространения для первого каскада и k -NN для второго) и алгоритм группировки цифр в числа позволил добиться уменьшения ошибки распознавания относительно существующих методов обработки топографических карт в 2,8 раз для полноты, в 2 раза для точности (таблица 2).

Экспериментальным способом была получена оптимальная последовательность операций итерационного алгоритма группировки ЛО. Было достигнуто уменьшение количества конечных точек на 30% по сравнению с современными разработками (таблица 3).

Таблица 2 – Сравнение результатов работы предложенного алгоритма распознавания ТО

Алгоритм	<i>Recall</i> , %	<i>Identity Precision</i> , %
<i>SLS-GHT</i> , 2017	–	93,3
<i>Faster R-CNN</i> , 2018	–	94,8
<i>Tesseract</i> , 2021	75,9	98,6
<i>Mask TextSpotter V3</i> , 2022	93,2	98,2
Предложенный	97,6	99,1

В процессе исследования алгоритма комплексного анализа было получено, что он позволяет полностью восстанавливать план горизонталей для текущего уровня ошибок распознавания чисел, при этом он может работать и в случае увеличения ошибки распознавания до уровня в 14% и вероятностью пропуска чисел до уровня в 25%

Разработанный алгоритм интерполяции сравнивался с методом линейной интерполяции, интерполяции кубическими сплайнами и интерполяцией Сибсона. Предложенный метод с фильтрацией даже при учёте 20% точек горизонталей имеет на 15% меньшее значение максимального абсолютного отклонения и на 15% меньшее значение среднеквадратичной ошибки.

Суммарное время автоматической обработки ЦТК можно оценить в 50 с/Мп, необходимое время на ручную обработку – 250 с/Мп, что значительно меньше текущего времени на векторизацию ЦТК в современных программных средах.

Пример преобразования ЦТК в ЦМР представлен на рис. 5, программный код разработанных функций и комплект преобразованных файлов для одной ЦТК доступен в репозитории *GitHub*: <https://github.com/Logiraz/analysis-and-processing-of-images-of-complex-structure>.

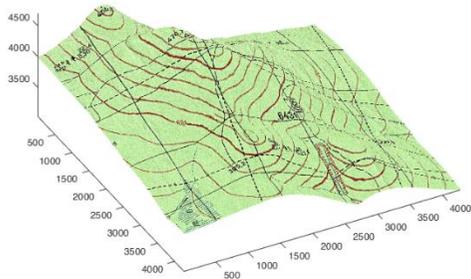
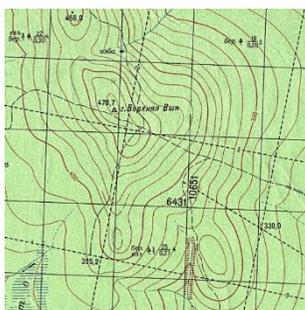


Рис. 5 – Результат работы алгоритмов преобразования ЦТК в ЦМР: слева – участок исходной ЦТК; справа – полученная ЦМР

В заключении проводится обобщение результатов проведённого диссертационного исследования.

Таблица 3 – Сравнение результатов работы предложенного алгоритма группировки ЛО

Алгоритм	Кол-во КТ на Мп	Кол-во ложных соединений на Мп
2006	294	6
2010	63	5
2012	325	6
2016	256	0
2018	46	
Предл.	31	6

ЗАКЛЮЧЕНИЕ

Основные результаты, которые были получены в ходе выполнения теоретических и экспериментальных исследований по теме, состоят в следующем.

– Проведён широкий и критический анализ методов и подходов, используемых при обработке цифровых изображений, анализе и распознаванию образов на них. Выявлены недостатки и преимущества оптического и интеллектуального подходов и было предложено использование гибридного, итерационного и комплексного подхода.

– Разработана классификация сложноструктурных изображений, основанная на способах и особенностях их получения, указаны их основные признаки. Для каждого типа описаны основные методы, использующиеся для их обработки.

– Разработана общая методика обработки сложноструктурных изображений, основанная на представлении их в виде точечных, линейных и площадных объектов, которую можно использовать для изображений различных типов приведённой классификации.

– Разработаны алгоритмы сегментации, локализации, распознавания, группировки и комплексного анализа объектов, находящихся на сложноструктурных изображениях. Для сегментации предложено использовать многометочный подход. Для локализации – поиск морфологических операций и операций фильтрации с помощью генетического алгоритма. Для распознавания – двухкаскадную систему, использующую для первого каскада метод скользящего окна, а для второго каскада – метод анализа суперсегментов. Для группировки точечных и линейных объектов приведены различные подходы к кластеризации, фильтрации и поиске отношений.

– Проведены экспериментальные исследования для нахождения оптимального типа и структуры интеллектуальных моделей, морфологических операций и операций фильтрации. В качестве исследуемых изображений выбраны смоделированные изображения морских навигационных карт и материалы спутниковых снимков. В обоих экспериментах была отработана методика обработки сложноструктурных изображений и показана возможность применения предлагаемых алгоритмов.

– Сформулированы ключевые проблемы преобразования изображения сканированных топографических карт в объёмную модель местности. Выявлены основные проблемы, возникающие при обработке топографических карт, как цветных сложноструктурных объектов низкой точности и высокой информативной плотности.

– Предложенные алгоритмы применены для сегментации, локализации, распознаванию и группировке объектов на изображениях топографических карт. Создан алгоритм восстановления плана горизонталей, состоящий из алгоритма соединения горизонталей и комплексного анализа распознанных значений на предыдущих этапах. На основе разработанного подхода к классификации интерполяционных областей был создан алгоритм, учитывающий характер интерполяции контурных линий в двумерную матрицу.

– Разработан программный комплекс, состоящий из нескольких модулей, реализующих разработанные алгоритмы, выполняющий преобразование топографической карты в объёмную модель местности с минимальным участием оператора.

Из указанных результатов следует вывод, что цель и задачи диссертационного исследования выполнены в полном объёме.

Перспективы развития настоящей темы исследования анализа и распознавания сложноструктурных изображений заключаются в следующих пунктах.

– Дальнейшая разработка алгоритмов, с помощью которых можно производить анализ и обработку над несколькими типами ССИ без изменения параметров и гиперпараметров интеллектуальных моделей и выполняемых операций.

– Разработка алгоритмов, учитывающих современное развитие аппаратного обеспечения, например, параллельных алгоритмов.

– Дальнейшее исследование стабильности указанных алгоритмов и интеллектуальных моделей и их оптимальности при решении задач распознавания.

– Создание онтологии обработки сложноструктурных изображений.

Публикации по теме диссертации

Статьи, опубликованные в рецензируемых научных журналах и изданиях, определённых ВАК РФ и Аттестационным советом УрФУ:

1. **Дмитриев Н.В.** Методика исследования устойчивости обучения нейронных сетей (на примере задачи классификации образов на сложноструктурных изображениях) / **Н.В. Дмитриев** // Научно-технический вестник Поволжья. – 2022. – № 12. – С. 108-111. (0,25 п.л.)

2. **Dmitriev N.V.** Complex method of reconstruction of contour lines / **Dmitriev N.V.** // AIP Conference Proceedings. proceedings of the 45th international conference on application of mathematics in engineering and economics (AMEE'19). – 2019. – p. 080005. (0,34 п.л.) (Scopus)

3. **Дмитриев Н.В.** Интерполяция распределённых данных горизонталей для получения цифровой модели рельефа / **Н.В. Дмитриев,**

В.С. Тарасян // Инженерный вестник Дона. – 2018. – №1 (48). – С. 85 (0,62 п.л. / 0,31 п.л.)

4. Тарасян В.С. Восстановление плана горизонталей при обработке топографических карт / В.С. Тарасян, **Н.В. Дмитриев** // Современная наука: актуальные проблемы теории и практики. Серия «Естественные и технические науки». – М.: Научные технологии. – 2017. – № 1. – С. 56-60. (0,44 п.л. / 0,22 п.л.)

5. Тарасян В.С. Сравнение методов сегментации пикселей топографических карт по типам объектов / В.С. Тарасян, **Н.В. Дмитриев** // Известия Тульского государственного университета: технические науки. – Тула: Тульский государственный университет. – 2016. – № 3. – С. 97-106. (0,62 п.л. / 0,31 п.л.)

Свидетельства о регистрации программы для ЭВМ:

6. Свидетельство о государственной регистрации программы для ЭВМ № 2019665260 RU. Программный комплекс преобразования изображений топографических карт в цифровую модель рельефа / **Дмитриев Н.В.** – № 2019665260; дата поступления 11.11.2019; дата государственной регистрации в реестре программ для ЭВМ 21.11.2019 (Российская Федерация).

7. Свидетельство о государственной регистрации программы для ЭВМ, № 2016611844 RU. Построение объёмной модели местности по топографической карте / Тарасян В.С., **Дмитриев Н.В.** – № 2016611844; дата поступления 18.12.2015; дата государственной регистрации в реестре программ для ЭВМ 11.02.2016 (Российская Федерация).

Другие публикации:

8. **Дмитриев Н. В.** Метод восстановления значений на топографических картах для создания цифровой модели рельефа / **Н.В. Дмитриев** // Прикладная информатика. 2019. – Т. 14. – № 6 (84). – С. 71-79. (1,0 п.л.)

9. Тарасян В.С. Гибридные процедуры автоматической обработки топографических карт / В.С. Тарасян, **Н.В. Дмитриев** // Лучшая научная статья 2017: сборник статей XII Международного научно-практического конкурса. – Пенза: МЦНС «Наука и Просвещение». – 2017. – С. 36-40. (0,31 п.л. / 0,15 п.л.)

10. **Дмитриев Н.В.** Комплексная система анализа сложноструктурных изображений / **Н.В. Дмитриев** // Интеграция образовательной, научной и воспитательной деятельности в организациях общего и профессионального образования: материалы IX Международной научно-практической конференции / отв. за выпуск Н.Ф. Сирина, д-р техн. наук. – Екатеринбург: УрГУПС, 2017. – Вып. 9 (228). – С. 194-198. (0,31 п.л.)

11. **Дмитриев Н. В.** Распознавание площадных элементов на топографических картах / **Н.В. Дмитриев**, В.С. Тарасян // Научно-

методический электронный журнал «Концепт». – 2017. – Т. 39. – С. 3476-3480. (0,81 п.л. / 0,40 п.л.)

12. **Дмитриев Н. В.** Алгоритм автоматического восстановления значений высот горизонталей / **Н.В. Дмитриев**, В.С. Тарасян // Молодежь в науке: Новые аргументы: Сборник научных работ V-го Международного молодежного конкурса. Часть I. – 2016. – С. 176-179. (0,43 п.л. / 0,22 п.л.)

13. **Дмитриев Н. В.** Автоматизированная система преобразования топографических карт в геоинформационную модель / **Н.В. Дмитриев** // Информационные технологии и нанотехнологии (конференция ИТНТ-2016). – Самара: Самарский научный центр РАН. – 2016. – С. 386-389. (0,25 п.л.)

14. **Дмитриев Н. В.** Автоматизированная система распознавания символов на топографических картах / **Н.В. Дмитриев**, В.С. Тарасян // Научно-методический электронный журнал «Концепт». – 2016. – Т. 11. – С. 1876-1880. (0,56 п.л. / 0,28 п.л.)

15. Тарасян В.С. Интеллектуальный комплекс 3D-визуализации топографических карт / В.С. Тарасян, **Н.В. Дмитриев** // Современные проблемы автоматизации и управления в энергетике и машиностроении: сборник трудов международной научно-практической конференции. – Пенза: Пензенский государственный технологический университет. – 2015. – С. 153-161. (0,56 п.л. / 0,28 п.л.)

16. **Дмитриев Н. В.** Система автоматизированного построения виртуальной модели местности по топографической карте / В.С. Тарасян, **Н.В. Дмитриев** // Молодежь в науке: Новые аргументы: Сборник научных работ II-го Международного молодежного конкурса. Часть I. – 2015. – С. 62-65. (0,25 п.л. / 0,12 п.л.)

17. Тарасян В.С. Интеллектуальная система анализа и преобразования топографических карт / В.С. Тарасян, **Н.В. Дмитриев** // Современные проблемы науки и образования. – 2015. – № 2-2. – С. 16. (0,5 п.л. / 0,25 п.л.)