

ОТЗЫВ

официального оппонента

Фаворской Маргариты Николаевны

на диссертацию Синявского Юрия Николаевича

«Непараметрические методы и программно-алгоритмический инструментарий для сегментации мультиспектральных спутниковых изображений», представленную на соискание ученой степени кандидата технических наук по специальности 05.13.18 – Математическое моделирование, численные методы и комплексы программ

Актуальность исследования

Предметом исследования диссертационной работы Синявского Юрия Николаевича являются непараметрические алгоритмы кластеризации на основе оценок плотности Розенблатта–Парзена (алгоритм MeanSC и ансамблевый алгоритм EMeanSC), для сегментации мультиспектральных изображений, полученных при помощи многозональной оптико-электронной спутниковой системы. Указанные алгоритмы позволяют разработать методы разделения формаций лесной растительности с близкими спектрально-яркостными характеристиками, а также реализовать оригинальный программно-алгоритмический инструментарий и стандартизованные веб-сервисы.

Проведенный соискателем анализ известных методов и алгоритмов обработки спутниковых изображений показал, что перспективным способом доступа к новым вычислительно эффективным алгоритмам кластеризации, применительно к задаче сегментации мультиспектральных спутниковых изображений, является построение распределенных геоинформационных систем, представленных в виде стандартизованных веб-интерфейсов. Это дает основание утверждать, что задачи, которым посвящена работа Синявского Ю.Н., являются актуальными.

Структура диссертационной работы

Диссертационная работа состоит из введения, пяти глав, заключения, списка литературы и трех приложений. Полный объем работы составляет 151 страницу. Список литературы содержит 154 наименования.

Во **введении** Синявский Ю.Н. обосновывает актуальность исследований, приводит степень разработанности темы, указывает цель и поставленные задачи, формулирует положения, выносимые на защиту, новые научные результаты, представляет практическую значимость и апробацию результатов.

В **первой главе** приводится классическая постановка задачи кластеризации и описание двух предшествующих ей задач – выбора меры схожести (или различия) объектов и выбора или выделения признаков. Раскрываются особенности задачи сегментации мультиспектральных спутниковых изображений. Построена классификация методов кластеризации с указанием конкретных методов, в основном разработанных за последние 30 лет. Подробно рассмотрены 5 групп непараметрических методов

(иерархические алгоритмы, методы разбиений, плотностные методы, сеточные методы и искусственные нейронные сети). Сделан вывод о необходимости комбинирования нескольких алгоритмов кластеризации или построении многоэтапных процедур, позволяющих на каждом этапе эффективно использовать достоинства отдельных алгоритмов.

Во **второй главе** приведена формальная постановка задачи кластеризации в рамках выбранного вероятностно-статистического подхода. Показано, что при использовании финитных радиально-симметричных ядер (для уменьшения вычислительной сложности) непараметрическая оценка Розенблатта–Парзена порождает вектор «среднего сдвига», сонаправленный с оценкой градиента плотности. Для сокращения трудоемкости процедуры «среднего сдвига» соискатель предложил альтернативную процедуру, основанную на сеточном подходе к кластеризации данных. Это позволило построить рабочую выборку, гарантированно содержащую представителей всех классов, избавиться от «шума» и многократно уменьшить вычислительную сложность алгоритма за счет сокращения числа анализируемых точек. В результате был разработан алгоритм MeanSC, использующий ядро Епанечникова. Для настройки алгоритма использовалось 3 параметра – порог отсекающего шума, параметр сглаживания и параметр объединения компонент связности. Исследование алгоритма проводилось как методом статистического моделирования на трех модельных наборах двумерных данных, так и на фрагментах реальных спутниковых снимков.

В **третьей главе**, посвященной разработке ансамблевого алгоритма кластеризации EMeanSC (Ensemble of MeanSC), обоснован выбор коллективного решения на основе согласованной матрицы различий кластеров. Показано, что ансамблевый алгоритм кластеризации EMeanSC позволяет улучшить качество результатов кластеризации по алгоритму MeanSC. Свойства алгоритма EMeanSC исследовались по методике исследования алгоритма MeanSC на модельных данных. Показано, что алгоритм EMeanSC способен успешно исправлять ошибки в случае сложных кластеров низкой плотности при наличии небольших расстояний между кластерами.

В **четвертой главе** приводится экспериментальное исследование предложенных алгоритмов на модельных данных, включая 7 наборов данных с разным количеством кластеров, формой кластеров, наличием пересечений и зашумленностью. При этом алгоритмы MeanSC и EMeanSC показали практически идентичные результаты по точности. Также экспериментальное исследование проводилось на 5 фотографиях (в цветовом пространстве RGB) и 3 спутниковых снимках (в цветовом пространстве RGB и Near InfraRed). Показано, что предложенные алгоритмы позволяют оперативно обрабатывать изображения высокого разрешения в диалоговом режиме.

В **пятой главе** представлена система веб-сервисов, доступная по протоколу WPS и включающая пять эффективных непараметрических алгоритмов, созданных в ИВТ СО РАН и реализованных в виде WPS-процессов. Также разработан пакет программ «Image Processing Toolkit» на

основе созданных и существующих алгоритмов сегментации спутниковых снимков. Приведено решение двух практических задач, а именно разделение формаций лесной растительности с близкими спектрально-яркостными характеристиками и обнаружение усыхающих кедровых древостоев по мультиспектральным изображениям высокого пространственного разрешения.

В **заключении** сформулированы основные выводы и результаты работы, показано, что цель диссертации достигнута путем решения поставленных задач.

В **Приложении 1** представлены графические пользовательские интерфейсы пакета «Image Processing Toolkit». **Приложение 2** содержит сканы 5 свидетельств о регистрации программ для ЭВМ по теме диссертационной работы. **Приложение 3** содержит скан акта об использовании результатов исследований.

Степень достоверности и обоснованности научных положений, выводов и рекомендаций

Целью диссертационной работы является разработка эффективных непараметрических алгоритмов сегментации спутниковых изображений и современной платформы для стандартизованного доступа к ним. Проанализирована отечественная и зарубежная литература по методам непараметрической кластеризации с учетом применимости к данной области. Синявский Ю.Н. корректно использует известные научные методы обоснования полученных результатов, выводов и рекомендаций. Представленные в работе положения и выводы обоснованы и аргументированы. Разработанные программные реализации предложенных алгоритмов и реализации существующих алгоритмов непараметрической кластеризации, включенные в пакеты для анализа данных с открытым исходным кодом, апробированы как на модельных наборах данных, так и на реальных спутниковых изображениях.

Результаты диссертационной работы неоднократно обсуждались на всероссийских и международных конференциях и опубликованы в 29 печатных работах, включая 7 статей, опубликованных в рецензируемых научных журналах, внесенных в перечень ВАК, 3 статьи, опубликованные в изданиях, индексируемых Scopus и Web of Science, 8 статей, опубликованных в других рецензируемых журналах, 4 доклада и 7 тезисов докладов международных и всероссийских конференций. Получены 5 свидетельств о государственной регистрации программ для ЭВМ.

Достоверность научных результатов подтверждается корректной постановкой задачи исследования, корректным применением методов теории вероятности и математической статистики, метода Монте–Карло и парадигм объектно-ориентированного программирования, непротиворечивостью полученных теоретических и практических результатов, апробацией полученных результатов, а также решением практических задач с помощью разработанного пакета программ «Image Processing Toolkit».

В целом, диссертационная работа изложена в грамотном, строгом научно-техническом стиле и оформлена в соответствии с требованиями ГОСТ. Автореферат диссертации полностью соответствует ее содержанию.

Научная новизна полученных результатов

В качестве новых научных результатов автором работы предложены:

– вычислительно эффективный непараметрический алгоритм кластеризации MeanSC на основе оценок плотности Розенблатта–Парзена для сегментации мультиспектральных спутниковых изображений. Эффективность достигается за счет введения сеточной структуры в пространстве признаков и перехода к рабочей выборке значительно меньшего объема. Сеточная структура впервые использована для повышения вычислительной эффективности поэлементного алгоритма кластеризации.

– подход к построению ансамбля непараметрических алгоритмов кластеризации, основанных на оценках плотности Розенблатта–Парзена, с помощью согласованной матрицы различий. На основе непараметрического алгоритма MeanSC создан ансамблевый алгоритм кластеризации EMeanSC, позволяющий обеспечить простоту настройки параметров и обработку мультиспектральных спутниковых изображений в диалоговом режиме.

– методы разделения формаций лесной растительности с близкими спектрально-яркостными характеристиками и обнаружения усыхающих древостоев по мультиспектральным изображениям. Методы позволяют обеспечить качественное выделение мелких и сильно пересекающихся кластеров.

Результаты в виде новых методов и алгоритмов, а также их интерпретация получены при активном участии автора. Проектирование и программная реализация разработанных алгоритмов, создание пакета программ «Image Processing Toolkit» и проведение численных экспериментов, а также реализация алгоритмов в виде модулей для GRASS GIS и в виде стандартизованных веб-сервисов, выполнены автором лично. Результаты полностью опубликованы в рецензируемых научных изданиях, прошли апробацию на конференциях различных уровней и использованы при создании картографических моделей в Институте почвоведения и агрохимии СО РАН, г. Новосибирск.

Значимость полученных результатов для науки и практики

Теоретическая значимость результатов диссертационного исследования определяется тем, что предложен вычислительно эффективный алгоритм кластеризации MeanSC для сегментации мультиспектральных изображений, разработанный на основе непараметрических оценок Розенблатта–Парзена с учетом особенностей спутниковых снимков. Также разработан непараметрический ансамблевый алгоритм кластеризации EMeanSC для сегментации мультиспектральных изображений, построенный на основе согласованной матрицы различий. Предложенные в диссертационной работе

решения вносят существенный вклад в развитие методов сегментации мультиспектральных изображений.

Практическая ценность диссертационной работы состоит в повышении эффективности сегментации спутниковых изображений в условиях малой априорной информации. Следует отметить, что разработанные алгоритмы превосходят описанные в литературе по качеству классификации и/или вычислительной эффективности. Алгоритмы внедрены в геоинформационную систему с открытым исходным кодом GRASS GIS, а также оформлены в виде стандартизованных веб-сервисов, что позволяет обеспечить доступ к ним по протоколу WPS.

Результаты диссертационной работы использованы при выполнении 3 междисциплинарных интеграционных проектов СО РАН (2003-2011 гг.), 6 проектов РФФИ (2009-2018 гг.), международного гранта фонда «Научный потенциал» («Human Capital Foundation», 2006 г.), а также гранта мэрии г. Новосибирска (2009-2010 гг.). По теме диссертационной работы получены 5 свидетельств о государственной регистрации программ для ЭВМ.

Разработанный программно-алгоритмический инструментарий передан в Институт почвоведения и агрохимии СО РАН для использования при крупномасштабном картографическом моделировании структурной организации растительности и почвенного покрова.

Выявленные недостатки и замечания

Изучение материалов диссертационной работы выявило следующие недостатки:

1. Представлены результаты кластеризации спутниковых изображений невысокого разрешения, а именно 550×500 пикселей и 1001×1045 пикселей (стр. 62). Не совсем понятно, каким образом (программно или с помощью эксперта) ассоциировались результаты кластеризации с названиями текстур на рис. 2.7 и 2.8 (стр. 61). Было бы целесообразно привести результаты кластеризации спутниковых снимков высокого разрешения с учетом оценок быстродействия алгоритма.

2. Следовало бы объяснить, почему разработанные алгоритмы MeanSC и EMeanSC показали практически идентичные оценки точности на модельных данных. Оценки точности алгоритмов для реальных изображений отсутствуют (глава 4). Также хотелось бы видеть оценки ошибок разработанных алгоритмов и известных алгоритмов.

3. Также следует отметить краткость выводов по всем пяти главам. Разделителем целой и дробной частей в числах является запятая, а не точка (для русскоязычных текстов). В тексте встречаются неудачные термины, например «взбирание по градиенту» (стр. 43).

Указанные замечания не влияют на общую положительную оценку работы.

**Заключение о соответствии диссертации требованиям и критериям,
установленным Положением о порядке присуждения ученых степеней**

Представленная к защите диссертационная работа Синявского Ю.Н. является завершенным научным исследованием, выполненным на актуальную тему, является целостной и логически обоснованной, содержит новые научные результаты и положения, а также рекомендации по их использованию на практике. Предлагаются новые вычислительно эффективные непараметрические алгоритмы кластеризации на основе оценок плотности Розенблатта–Парзена для сегментации мультиспектральных спутниковых изображений (MeanSC и ансамблевый алгоритм EMeanSC). Разработан пакет программ «Image Processing Toolkit», позволяющий выполнять визуализацию, анализ и обработку спутниковых изображений.

Оформление диссертации и автореферата удовлетворяет требованиям соответствующих нормативно-методических документов. В автореферате диссертационной работы полностью отражены основные полученные результаты и дано краткое изложение содержания выполненных исследований. Содержание разделов автореферата полностью соответствует содержанию диссертационной работы.

Представленная диссертационная работа соответствует требованиям п. 9 «Положения о порядке присуждения ученых степеней» постановления Правительства Российской Федерации от 24.09.2013 г. № 842, предъявляемым к диссертациям на соискание ученой степени кандидата наук, а ее автор Синявский Юрий Николаевич заслуживает присуждения ему ученой степени кандидата технических наук по специальности 05.13.18 – Математическое моделирование, численные методы и комплексы программ.

Официальный оппонент

Доктор технических наук, профессор,
заведующий кафедрой информатики
и вычислительной техники ФГБОУ ВО
«Сибирский государственный университет
науки и технологий имени
академика М.Ф. Решетнева»



Фаворская Маргарита Николаевна

19 мая 2021 г.

Адрес места работы:

660037, г. Красноярск, проспект им. газеты Красноярский рабочий, д. 31

<http://www.sibsau.ru/>

Тел. 8 (391) 213-96-23

e-mail: favorskaya@sibsau.ru

