

Моделирование рассуждений при поиске и описании объектов на космоснимках

Д.Р. Касимов¹

¹ Ижевский государственный технический университет имени М.Т. Калашиникова, 426069, Россия, г. Ижевск, ул. Студенческая, д. 7

Аннотация

В статье представлен подход к проблеме контекстного поиска и описания объектов на растровых космоснимках, заключающийся в моделировании рассуждений на основе структурированных прецедентов. В результате обработки изображения строится граф смежности цветовых областей. Объект характеризуется цветом, атрибутами формы отрезков границы и формы объекта в целом. Структурированный прецедент представляется в виде лучевого графа, дуги которого упорядочены в соответствии с положительным обходом границ областей. С помощью алгоритма сопоставления графов в анализируемом изображении выявляются вхождения прецедентов из базы данных системы. При обнаружении вхождения применяется правило прецедентного вывода. Степень принадлежности объекта некоторому классу зависит не только от свойств самого объекта, но и от достоверности окружающих его объектов. Стратегия контекстного поиска содержит этапы рекурсии и итерации. В отличие от нейросетевых технологий, предложенный подход позволяет не только классифицировать изображенные объекты, но и получать их структурированные описания. Кроме того, выдаваемое системой классификационное решение имеет аргументированное обоснование. Результаты эксперимента показывают, что рассуждения на основе структурированных прецедентов позволяют уточнять результаты классификации и повышать достоверность распознавания объектов на космоснимках.

Ключевые слова: компьютерное зрение, цифровая обработка изображений, распознавание образов, структурный анализ, сегментация, аппроксимация, граф смежности, лучевой граф, рассуждения по прецедентам.

Цитирование: Касимов, Д.Р. Моделирование рассуждений при поиске и описании объектов на космоснимках / Д.Р. Касимов // Компьютерная оптика. – 2020. – Т. 44, № 5. – С. 772-781. – DOI: 10.18287/2412-6179-CO-716.

Citation: Kasimov DR. Searching and describing objects in satellite images on the base of modeling reasoning. Computer Optics 2020; 44(5): 772-781. DOI: 10.18287/2412-6179-CO-716.

Введение

Задача обнаружения и описания объектов заданных классов на спутниковых снимках особенно актуальна в связи с бурным ростом объема подобной информации. Проблема заключается в том, что исходные данные слабо структурированы – изображения представлены в виде множества разноцветных пикселей, а требуется получить (обнаружить) объекты, принадлежащие указанным классам, и описать их, т.е. вычислить признаки – факты, на основании которых найденный объект отнесен к тому или иному классу.

Автоматическая классификация объектов на космоснимках является сложной и крайне актуальной проблемой в области компьютерного зрения, привлекающей внимание многих исследователей. Некоторые из разработанных техник идентифицируют объекты с высокой точностью, приемлемой для ряда практических целей, но пока лишь в тех конкретных условиях (разрешение изображения, тип местности, время года), под которые проектировались. В современных исследованиях акцент начал смещаться в сторону создания моделей и методов идентификации, обладающих таким уровнем обобщающей способно-

сти [1], который позволит значительно расширить сферу их практической применимости.

Для обнаружения заданных объектов на изображениях широко используется подход, основанный на искусственных нейронных сетях (ИНС). В рамках этого подхода идет постоянное совершенствование методов обучения ИНС, создаются многослойные, гибридные (в различных комбинациях), сверточные нейронные сети. Например, в работе [2] применяется серия нейросетевых детекторов, один из которых ориентирован на выделение дорог, а все остальные направлены на распознавание зданий определенного размера. В работе [3] классификацию пикселей спутникового снимка производит сеть с глубокой архитектурой SegNet [4], обладающая деконволюционными слоями. В работе [5] решается задача выделения дорог с помощью нейронной сети и ряда приемов постобработки производимых ею результатов (векторизация, контроль ширины, устранение разрывов и др.). Одной из ключевых проблем обучения ИНС является разметка тестовых данных и поддержка их актуальности [6]. В случае, если ИНС, обученная на снимках местности одного типа, применяется к снимкам местности другого типа, ее эффективность заметно снижается.

Другой класс методов идентификации объектов основывается на вычислении интегрированных спектральных характеристик участков изображений. Например, в [7–9] определение зданий на изображениях осуществляется на основе так называемого морфологического индекса здания (МВИ), предполагающего выполнение морфологического преобразования "white top-hat" и вычисление дифференциальных морфологических профилей. В ходе обработки, как и в случае применения ИНС, попиксельно вычисляется карта бинаризации, отображающая результаты классификации. Подобного рода методы лежат в плоскости дескриптивного подхода [10] к анализу изображений. В дескриптивном подходе алгоритм анализа изображения представляется в форме алгебраического выражения, элементами которого являются модели изображения и операции их обработки (трансформации). В качестве описания области изображения служит, как правило, n -мерный признаковый вектор (дескриптор). Например, в работе [11] описанием изображения служит поле градиентов – множество двумерных векторов, представляющих горизонтальные и вертикальные производные яркости пикселей изображения; на основе поля градиентов алгебраически вычисляют доминирующую ориентацию текстуры, определяют степень размытия изображения, оценивают сходство двух изображений, осуществляют поиск объекта на аэроснимке по заданному изображению-образцу и др. В работе [12] производится детальный анализ различных методов выявления особых точек (локальных особенностей) на изображении, их описания и сопоставления на основе дескрипторов. В рамках дескриптивного подхода объекты распознаются как целостные образы, а не как составные пространственные конфигурации. Опираясь на алгебраические принципы, весьма проблематично формировать богатые семантические описания, достаточно полно и выразительно характеризующие содержание изображений (состав и форма объекта, его взаимосвязи с другими объектами, ситуации).

Широко известным подходом к анализу изображений является синтаксический [13] (другие названия – структурный, лингвистический), в котором для целевого класса объектов (в общем случае, сцен) создается структурное описание в виде формальной грамматики, а определение того, принадлежит ли предъявленный объект целевому классу, происходит путем выполнения грамматического разбора. Применяя такой подход в чистом виде, можно рассчитывать на успешное распознавание лишь тех объектов, синтаксическая структура которых весьма точно соответствует правилам грамматики. Поскольку космоснимкам присуща значительная вариативность объектов и их сочетаний, проблематично привести примеры успешного решения задач анализа космоснимков на основе синтаксического подхода.

Следует отметить, что оценочные метрики, применяемые в работах, посвященных проблеме обнару-

жения объектов на космоснимках, как правило, работают на пиксельном уровне. На взгляд автора, в ряде прикладных задач поддержки принятия решений по анализу космоснимков важно предоставить пользователю не просто набор выявленных объектов, но также описания этих объектов и аргументированное обоснование причины отнесения каждого объекта к тому или иному классу.

Целью настоящей работы является повышение релевантности результатов структурного анализа изображений для поддержки принятия решений. Ставится задача автоматического обнаружения и оконтуривания строений на аэрокосмических снимках, которая имеет большое практическое значение, например, для реализации оперативного дистанционного контроля застройки и землепользования.

Предлагаемый подход заключается в моделировании рассуждений на основе структурированных прецедентов. Под моделированием рассуждений в системах искусственного интеллекта понимается логический вывод – получение новых фактов из данных с помощью аксиом и правил. Суть предлагаемого решения заключается в вычислении информативных признаков с помощью методов цифровой обработки изображений и двухэтапной классификации объектов: вначале с помощью решающих правил на основе полученных признаков, затем – для оставшихся неклассифицированными объектов – на основе прецедентов.

Вывод на основе прецедентов успешно применяется в интеллектуальных системах поддержки принятия решений. Но применение этого подхода к анализу изображений затрудняется некоторыми их специфическими особенностями, в частности, такими как неструктурированность исходных данных и зависимость семантики от контекста – пространственных отношений между элементами объектов.

Чтобы решить указанную проблему, потребовалось разработать методы выделения признаков, характеризующих форму объектов и отношения между ними, структурирования и реструктурирования объектов, поиска прецедентов и вычисления степени сходства описаний пространственных ситуаций. В целом, данная работа вносит вклад в дальнейшее развитие методов и инструментов вышеупомянутого структурного подхода применительно к задаче анализа космоснимков.

1. Предлагаемый подход

Ниже представлена общая стратегия анализа изображения:

1. Разбиение изображения на кадры (части) равного размера с целью повышения качества последующей цветовой сегментации за счет большей локальности. На полном большом снимке нередко можно увидеть разные типы местности (лес, поле, реку, жилой массив), что сильно затрудняет улавливание слабо выраженных цветовых различий между объектами. Чем меньше размер анализиру-

емого участка изображения, тем лучше объекты отделяются друг от друга в ходе цветовой сегментации. Это позволяет уменьшить количество цветных кластеров и, соответственно, объем вычислений.

2. Обработка каждого кадра с целью выделения цветных областей и аппроксимации их границ:

- 1) Цветовая сегментация.
- 2) Выделение границ цветных сегментов.
- 3) Аппроксимация границ.

Применяемые здесь методы подробно рассмотрены в работе [14]. Цветовая сегментация основана на построении трехмерной гистограммы в цветовом пространстве HSV и поиске локальных максимумов путем сканирования пространства трехмерным оператором анализа окрестности. Чтобы ускорить процесс и снизить излишнюю детализацию результатов (избежать пересегментации), все цветовое пространство предварительно квантуется на относительно небольшое число цветных диапазонов. Экспериментальным образом была подобрана следующая оптимальная степень «огрубления» цветового пространства HSV: $15 \times 7 \times 15$. В рамках кадра такой уровень аппроксимации позволяет достаточно хорошо разделять разные объекты друг от друга (по тону, по насыщенности, по яркости). На рис. 1 приведен пример цветовой сегментации отдельного кадра изображения и выделения границ полученных цветных сегментов. В результате кластеризации цветной палитры первоначально сформировалось 33 цветных сегмента, число которых затем было сокращено до 7 наиболее значимых с целью устранения излишних деталей и упрощения последующего анализа.

3. Слияние цветных областей из соседних кадров по переходящим отрезкам границ. Цветные области, относящиеся к одному объекту и лежащие с противоположных сторон от границы соседних кадров, как правило, обладают смежными участками границы (горизонтальными либо вертикальными отрезками) и схожим цветом. Такие области можно выявить и объединить, тем самым снизив степень фрагментации объектов, расположенных одновременно в двух или более смежных кадрах. Эксперименты показывают, что для строений ошибки слияния происходят крайне редко. Это можно объяснить тем, что области строений имеют малую степень извилистости и высокую степень сходства по цвету.

Если случается, что объединяются области, относящиеся к разным объектам, остается шанс их отделения друг от друга на следующем шаге.

4. Разбиение цветных сегментов на подобласти в узких «перешейках» с целью отделения объектов разных классов друг от друга. Это целесообразно делать в связи с тем, что такого рода «перешейки», как правило, не присущи реальным объектам.

5. Формирование графа смежности областей.

6. Вычисление и фазификация значений атрибутов областей. Каждый объект характеризуется геометрической формой, цветом, отношениями смежности с соседями. Характер общей границы, а именно форма и протяженность, является полезной информацией для принятия решения о классах смежных объектов. Например, общая граница между домом и тенью, как правило, является прямолинейной, а между домом и деревом – извилистой.

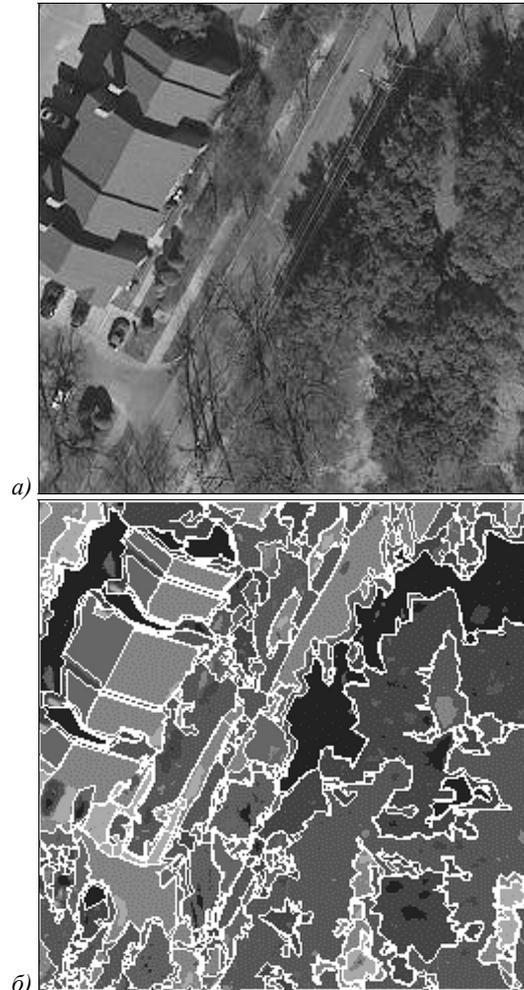


Рис. 1. Пример цветовой сегментации и выделения границ: а) исходный кадр; б) результирующие цветные области и их границы

7. Классификация объектов на основе признаков. Объект может быть отнесен одновременно к нескольким целевым классам с разной степенью принадлежности.

8. Вывод по прецедентам – рекурсивное уточнение соседей ранее классифицированных объектов путем сопоставления графов ситуации и прецедента и вывода классификационного решения на его основе.

В ходе обработки из изображения извлекается следующая информация:

$IMAGE = \langle v0, FM, REG, EDGE, A, HAR, HIR, GA \rangle$, где $v0$ – корневая вершина с общими атрибутами изображения;

FM – матрица фреймов (кадров);
 REG – множество областей;
 $EDGE$ – множество графов границ областей – циклических списков дуг и отрезков;
 A – множество атрибутов;
 HAR – гиперграф отношений смежности (adjacency) областей;
 HIR – гиперграф отношений вложенности (insert) областей;
 GA – двудольный граф атрибутов.

Гиперграф отношений смежности областей:

$$HAR = H(REG, ER),$$

где ER – множество отношений смежности между отрезками границ областей, $er = \langle fm(i1, j1), reg(k1), e(l1); fm(i2, j2), reg(k2), e(l2) \rangle$.

Гиперграф отношений вложенности областей:

$$HIR = H(REG, RR), rr = \langle reg1, reg2 \rangle.$$

Двудольный граф атрибутов:

$$GA = G(V, A), V = v0 \cup FM \cup REG \cup E.$$

Ниже приведены геометрические и цветовые признаки объектов. Значения признаков переводятся из количественного представления в качественное нечеткое, в форму лингвистических переменных [15]:

- *Прямолинейность* (ОченьМалая, Малая, Средняя, Высокая, ОченьВысокая) – отношение суммы длин значимых (достаточно длинных) линий к периметру границы.
- *Доля значимых прямых углов* (ОченьМалая, Малая, Средняя, Большая, ОченьБольшая) – отношение числа прямых углов, образованных значимыми линиями, к общему числу углов.
- *Наличие трех сторон прямоугольника* (Да, Нет). На рис. 2 показаны примеры областей, обладающих данным признаком.



Рис. 2. Области, обладающие признаком наличия трех сторон прямоугольника (отображены белым)

- *Наличие двух значимых параллельных линий* (Да, Нет).
- *Наличие одной значимой линии* (Да, Нет).
- *СредняяШирина* (ОченьМалая, Малая, Средняя, Большая) – отношение средней длины сечения к длине скелетона.

- *Удлиненность* (Малая, Средняя, Большая, ОченьБольшая) – соотношение сторон минимального описывающего прямоугольника.
- *Площадная прямоугольность* (Малая, Средняя, Большая, ОченьБольшая) – отношение точной площади объекта к площади минимального описывающего прямоугольника.
- *Извилистость* (Малая, Средняя, Высокая, ОченьВысокая) – число смен знака приращения угла при обходе элементов границы.
- *Светлота* (ОченьМалая, Малая, Средняя, Высокая, ОченьВысокая).
- *ЦветЗелени* (Да, Нет), *ЦветПожелтевшейТравы*, *ЦветАсфальта*, *ЦветБетона*, *ЦветСпортивно-гоПокрытия* и т.д.

Функции принадлежности (μ) нечетких значений лингвистических переменных имеют трапецидальную форму.

Классификация объектов осуществляется в два этапа. На первом этапе с помощью описаний, примеры которых представлены в табл. 1, устанавливается степень принадлежности каждого объекта к целевым классам.

Второй этап классификации объектов осуществляет анализ совокупностей объектов в соответствии с прецедентами. На снимках высокого разрешения объектам типа строений и деревьев всегда сопутствуют тени, падающие в направлении, противоположном положению солнца. В частности, этот факт был использован в работе [9] для уточнения результатов классификации на этапе постобработки, что говорит о том, что даже простейшие прецеденты могут приносить ощутимую пользу. В целом, выявление известных ситуаций на изображениях с помощью прецедентов способствует более релевантному их описанию.

Для представления прецедентов используется нечеткий лучевой граф [16] – ориентированный граф типа дерево, который описывает объекты изображения и их пространственное положение относительно корневой вершины с помощью нечетких атрибутов.

На рис. 3 представлены примеры прецедентов.

Дуги лучевого графа отображают пары объектов, связанных отношением смежности. Они упорядочены в соответствии с положительным обходом границы области опорного объекта. В прецеденте объектам (не обязательно всем) приписываются метки целевых классов. Также на конкретный объект может быть наложен набор требований к значениям признаков, в т.ч. в виде диапазонов допустимых значений. Прецеденты создаются экспертом и хранятся в базе данных системы. Прецеденты могут иметь разную обобщающую способность, варьируясь от типовых до редко встречающихся случаев, что зависит от числа узлов и строгости требований в каждом узле.

Используемая модель представления прецедентов удобна тем, что описание строится, отталкиваясь от заданного опорного объекта. При этом четко выра-

жаются намерения автора прецедента (эксперта). Данный принцип описания структур хорошо себя зарекомендовал в дескриптивных логиках [17] – семей-

стве языков представления знаний, сочетающих богатую выразительность и приемлемую вычислительную сложность решения проблем.

Табл. 1. Примеры описаний классов объектов

Класс объекта	Общие требования	Разновидности	Степень принадлежности
Здание	(Удлиненность ≤ Малая) ∧ (СредняяШирина ≥ Большая) ∧ (ЦветЗелени = Нет) ∧ (Светлота ≥ Средняя)	(Прямолинейность ≥ ОченьВысокая)	1,0
		(Прямолинейность ≥ Высокая) ∧ (ДоляПрямыхУглов ≥ Большая) ∧ (Извилистость ≤ Малая) ∧ (ПлощаднаяПрямоугольность ≥ Средняя)	1,0
		(ПлощаднаяПрямоугольность ≥ Высокая) ∧ (ДоляПрямыхУглов ≥ Большая)	1,0
Тень от искусств. объекта	(Светлота ≤ Малая) ∧ ((ЦветЗелени = Нет) ∨ (Светлота = ОченьМалая))	(Прямолинейность ≥ ОченьВысокая)	1,0
		(Прямолинейность ≥ Высокая)	0,8
		(Прямолинейность ≥ Средняя)	0,5
Фрагмент дороги	(Прямолинейность ≥ Высокая) ∧ ((Удлиненность ≥ Большая) ∨ (СредняяШирина ≤ Малая)) ∧ ((ЦветАсфальта = Да) ∨ (ЦветБетона = Да))	-	1,0

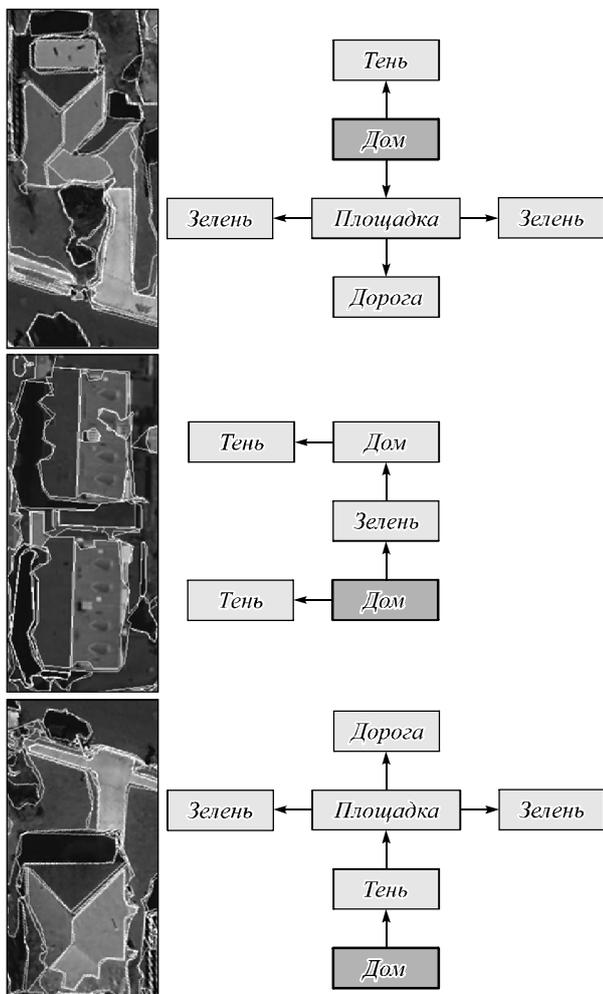


Рис. 3. Примеры прецедентов

С помощью алгоритма сопоставления графов, например, [18–19], в анализируемом изображении

выявляются вхождения прецедентов из базы данных системы. На вход алгоритма сопоставления подаются граф изображения и граф прецедента. В графе изображения алгоритм выявляет подграфы, изоморфные графу прецедента. Требование изоморфизма является не строгим (задача класса inexact subgraph matching). Процесс сопоставления начинается с корневой вершины лучевого графа прецедента и продолжается рекурсивно по ребрам, исходящим из совпавших на текущем шаге вершин. Для каждого найденного вхождения алгоритм выдает соответствия между узлами и ребрами графов, а также общую оценку сходства ($\in [0, 1]$) графа прецедента с найденным подграфом изображения. Значение общей оценки сходства вычисляется путем интеграции следующих факторов:

- доля узлов и ребер графа прецедента, которым найдено соответствие в графе изображения;
- величина отклонения значений атрибутов узлов и ребер графа изображения от требований, заданных в соответствующих узлах и ребрах графа прецедента;
- степень принадлежности объекта изображения к целевому классу, указанному в прецеденте.

На рис. 3 можно заметить, что строение в графе прецедента представлено одной вершиной, а на изображении оно состоит из нескольких цветовых областей. Чтобы было возможно установить соответствие, к анализируемому изображению применяется следующий алгоритм сборки единого объекта из нескольких цветовых областей:

1. Выборка областей-кандидатов по набору критериев (по величине, по форме – прямоугольные, треугольные, трапециевидные).
2. Определение среди кандидатов смежных групп путем выявления связанных компонент в графе смежности.

3. В каждой смежной группе выбор варианта подгруппы областей, имеющих схожие оттенки.
4. Синтез невыпуклой оболочки вокруг выбранной подгруппы областей, вычисление признаков созданной суперобласти.
5. Подстановка созданной суперобласти под вершину графа прецедента; если оценка сходства недостаточная, то рассмотреть другой вариант подгруппы.

Следует отметить, что среди прецедентов существуют также такие, в которых каждая часть строения представлена отдельной вершиной графа. Соответственно, они не требуют применения вышеприведенного алгоритма. Однако представление строения в виде единого цельного объекта является предпочтительным, так как в этом случае в прецеденте на форму каждого отдельного фрагмента не накладываются жесткие ограничения. Это очень важно в условиях проблемы искаженности и неточности границ областей, формируемых процедурой автоматической цветовой сегментации изображения.

Когда в изображении обнаруживается вхождение прецедента с достаточно высокой оценкой сходства, информация о классах объектов, содержащаяся в прецеденте, переносится соответствующим объектам изображения.

Правило прецедентного вывода степени принадлежности объекта известному классу выглядит следующим образом:

$$IF a \in C \wedge Correspond(a, b) \wedge b \underset{\mu < \lambda}{\in} C THEN b \underset{\mu = \lambda}{\in} C,$$

где a – объект из прецедента; C – метка целевого класса, назначенная объекту в прецеденте; b – объект изображения; *Correspond* – отношение соответствия объектов, установленное алгоритмом сопоставления графов; $\underset{\mu}{\in}$ – отношение (нечеткой) принадлежности объекта целевому классу со степенью μ ; λ – выведенная степень принадлежности.

Значение λ вычисляется по эмпирической формуле, учитывающей общую оценку сходства графов, а также надежность классификации соседних объектов (соседи влияют на значение с большим весом).

Процесс прецедентного вывода имеет рекурсивный характер. Классифицированные на текущем шаге объекты становятся опорными для следующего шага. Применяется стратегия поиска в ширину. Процесс повторяется, пока добавляется новая информация.

2. Эксперименты

Экспериментальное исследование представленного подхода проведено на снимках, взятых из набора данных Inria Aerial Image Labeling Dataset [1]. Изображения являются цветными, имеют размер 5000×5000 с пространственным разрешением 30 см (сверхвысокое разрешение). Набор данных [1] содержит также эталонную пиксельную разметку зданий, предназначенную для объективной оценки качества детекции зданий.

Каждое исходное изображение разбивалось на кадры размером 500×500 пикселей. После этапа покрупной обработки отдельные кадры «сшивались» обратно в единое изображение исходного размера.

Рис. 4 предоставляет читателю возможность визуально оценить эффект от применения алгоритма прецедентного вывода, а также сравнить результаты работы нашего подхода с результатами работы двух представителей класса нейросетевых методов.

Видно, что на основе прецедентного вывода качество выделения строений повышается. Можно заметить, что, несмотря на наличие промежутков между близко расположенными друг к другу домами, альтернативный метод FCN нередко выделяет их как одно большое строение. В нашем случае эта ошибка встречается заметно реже. Можно также наблюдать, что границы строений, которые выдает нейросетевой метод, являются достаточно сильно сглаженными и не всегда соответствуют действительности. В то время как в результатах работы предложенного подхода границы всех выделенных объектов прямолинейные, большое число их углов и сторон весьма точно соответствуют реальным строениям.

Количественная оценка эффективности подхода произведена на множестве изображений, состоящем из 15 снимков территории West Tyrol (Европа). Данный набор снимков суммарно покрывает площадь 33,75 км² и содержит порядка 8 тыс. строений.

В табл. 2 представлены полученные экспериментальные оценки эффективности системы, вычисленные на уровне пикселей. Пиксельная полнота – сколько (%) пикселей среди тех, что относятся к строениям в эталонной разметке, были успешно обнаружены системой. Пиксельная точность – сколько пикселей среди тех, что были отнесены системой к строениям, действительно являются пикселями строений. Показатель *IoU* вычисляется по формуле:

$$IoU = |D \cap E| / |D \cup E|,$$

где D – пиксели изображения, помеченные системой как пиксели строений; E – пиксели строений в эталонной разметке.

Показатель *IoU* является комплексной характеристикой качества семантической сегментации, и именно на него обращают основное внимание в работах по семантической сегментации изображений [20].

В табл. 3 представлены показатели, вычисленные на уровне объектов. Объектная полнота – какую долю целевых объектов удалось обнаружить. Объектная точность – сколько выданных системой объектов являются верными. Значение *F1*-меры вычисляется следующим образом:

$$F1 = 2 \times (R \times P) / (R + P),$$

где R – объектная полнота, P – объектная точность.

Объект считался верно классифицированным как строение, если он пересекался с целевым строением в эталонной разметке минимум на 50 % площади.

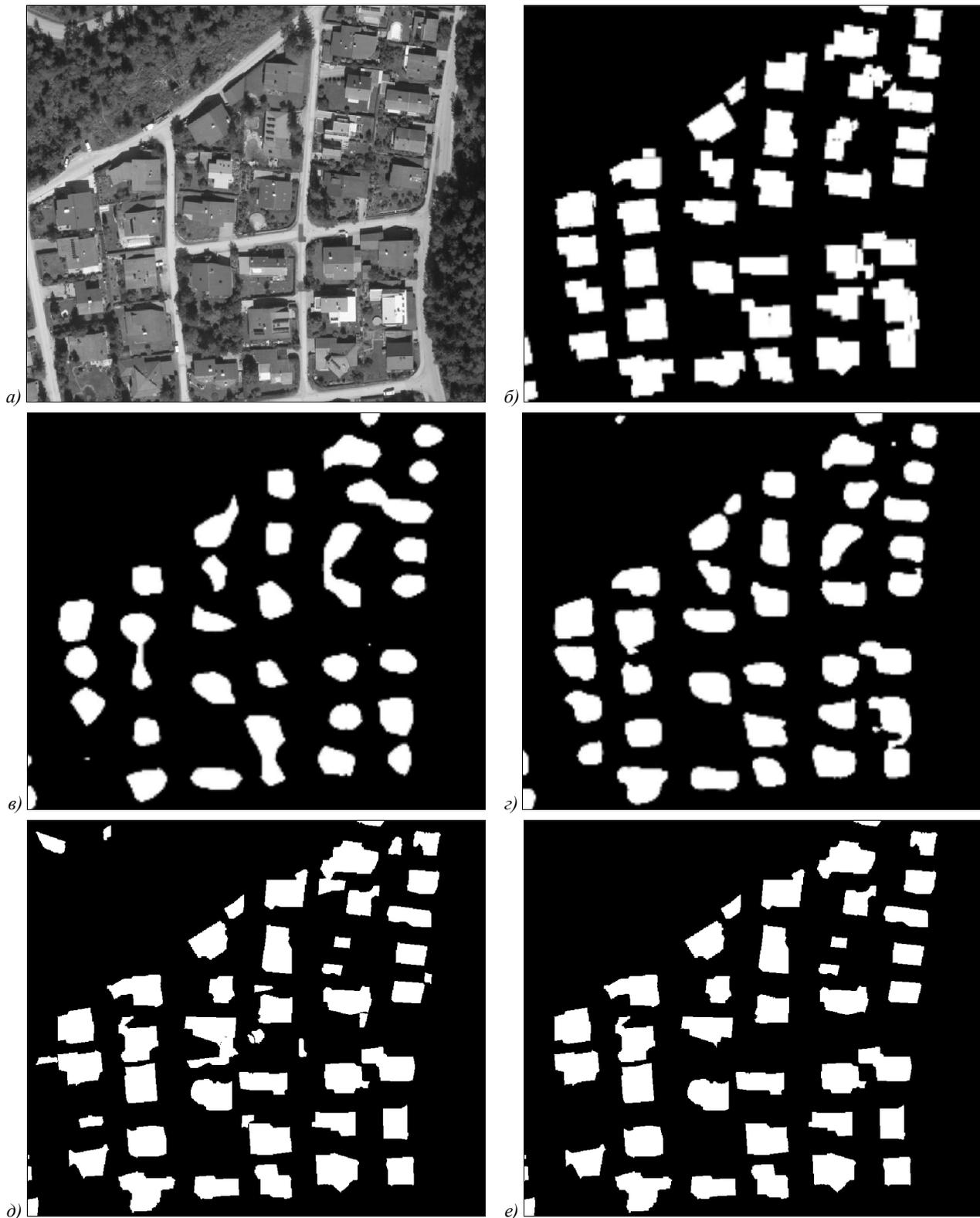


Рис. 4. Сравнительные результаты выделения строений на аэрофотоснимке: а) фрагмент исходного изображения; б) эталонная пиксельная разметка строений; в) результат работы нейронной сети FCN [1]; г) результат работы нейронной сети MLP [1]; д) результат работы представленной системы без использования прецедентов; е) результат работы представленной системы с применением прецедентного уточнения

Предложенный алгоритм прецедентного вывода позволил повысить общее качество детекции строений на 9,34% в случае пиксельно-ориентированного

способа оценки и на 45,53% в случае объектно-ориентированного метода оценки. Большее изменение во втором случае связано с тем, что в ходе преце-

дентного вывода устраняется много ложных мелких объектов, которые влияют на объектные показатели гораздо сильнее, чем на пиксельные.

Нейросетевой метод FCN на этом же наборе данных дает $IoU = 46,86$ [1]. Представленная система работает на 5,68% лучше.

Табл. 2. Оценки качества детекции строений, вычисленные на уровне пикселей

№ снимка	Без применения прецедентов			С применением прецедентов		
	Полнота	Точность	IoU	Полнота	Точность	IoU
1	56,40	63,44	42,57	56,58	77,73	48,68
2	65,56	46,28	37,23	63,59	67,62	48,75
3	60,92	63,86	45,30	58,39	75,10	48,92
4	69,24	52,36	42,48	70,22	70,73	54,41
5	59,85	58,65	42,09	58,19	71,84	47,38
6	68,07	52,55	42,16	68,33	70,71	53,26
7	67,87	49,28	39,96	67,40	73,18	54,05
8	69,98	62,83	49,49	70,93	77,72	58,94
9	63,33	72,94	51,28	62,25	79,31	53,54
10	64,46	64,37	47,51	63,44	77,68	53,66
11	66,33	57,34	44,42	65,96	71,48	52,22
12	65,60	40,07	33,11	63,83	62,73	46,28
13	75,42	35,53	31,84	74,24	63,38	51,96
14	70,69	56,83	46,00	68,44	71,55	53,80
15	72,17	65,82	52,50	71,72	81,01	62,30
	66,39	56,14	43,20	65,57	72,78	52,54

Табл. 3. Оценки качества детекции строений, вычисленные на уровне объектов

№ снимка	Без применения прецедентов			С применением прецедентов		
	Полнота	Точность	F1-мера	Полнота	Точность	F1-мера
1	85,14	10,32	18,40	76,25	63,71	69,42
2	81,13	13,10	22,56	70,42	50,61	58,89
3	85,19	15,23	25,84	76,70	66,05	70,98
4	85,27	08,34	15,20	74,93	47,97	58,49
5	87,96	10,13	18,17	82,20	57,30	67,53
6	86,39	09,83	17,65	78,70	53,52	63,71
7	86,02	06,81	12,63	77,42	47,16	58,62
8	87,68	20,60	33,36	83,03	71,73	76,97
9	94,66	21,72	35,33	90,29	68,43	77,86
10	88,34	17,03	28,55	82,22	71,19	76,31
11	84,56	11,21	19,79	78,23	58,86	67,17
12	88,20	05,09	09,63	78,65	39,44	52,53
13	88,14	02,53	04,92	80,51	32,53	46,34
14	85,91	09,30	16,79	80,11	51,06	62,37
15	100,00	15,59	27,04	97,12	70,31	81,57
	87,64	11,79	20,39	80,45	56,66	65,92

Двум другим нейросетевым методам из работы [1] – MLP и Skip – представленный подход несколько проигрывает. Превосходство метода MLP составляет 5,41%, а метода Skip – 2,37%.

К сожалению, альтернативные методы не были оценены авторами по объектным показателям, которые для технологий класса GEOBIA (Geographic Object-based Image Analysis – объектно-ориентированного анализа географических изображений) [21] являются более важными.

Одной из причин снижения показателей является наличие узких полос на границе строений и теней, из-

за чего страдает пиксельная точность результата. В целом, результат сильно зависит от качества цветовой сегментации. На одном из снимков было оценено, какой максимальный показатель качества работы гипотетически можно достичь на основе цветowych сегментов, формируемых процедурой сегментации. Для этого с помощью эталонной разметки была произведена выборка цветowych областей, у которых не менее 80% пикселей помечены в эталонной разметке как пиксели строений. Далее для множества выбранных областей были вычислены показатели релевантности на уровне пикселей: $Полнота = 79,91$, $Точность =$

91,16, $IoU=74,16$ (при повышении порога минимального пересечения значение IoU становится еще меньше). Таким образом, подтверждено, что неточность границ цветовых сегментов оказывает серьезное влияние на качество оконтуривания строений. Разработка методов, позволяющих уменьшить данное влияние, например, техник локального анализа для уточнения границ выявленных зданий, является предметом будущих исследований.

Следует отметить, что, в отличие от альтернативных методов, в предложенном подходе в качестве результата выдаются не только обнаруженные объекты, но также и их содержательное описание, что очень важно для систем поддержки принятия решений. Форма описания объектов показана в табл. 1.

Нейронные сети по качеству обнаружения строений остаются несколько впереди. Однако представленный подход особенно востребован в тех случаях, когда по определенным причинам применение нейронных сетей является вовсе неприемлемым, например, по причине того, что они не предоставляют аргументированное обоснование своих решений (в отличие от представленного подхода).

Заключение

В предложенном подходе к поиску и описанию объектов на изображениях, основанном на моделировании рассуждений на пространственных данных, принятие решения о классе объекта происходит на основе определения формы отрезков границы и формы объекта в целом, а также соседства объектов.

На реальных снимках одна и та же форма может принимать различные интерпретации при разном контексте (окружении), и здесь в предложенном подходе свою роль начинает играть граф смежности областей.

В отличие от популярных в настоящее время технологий, применяющих нейронные сети, интегрированные спектральные дескрипторы, предложенный подход позволяет не только классифицировать изображенные объекты, но и получать их структурированные описания. Формируемые результаты содержат описания и их обоснования, легко интерпретируемые людьми.

В подходе не требуется длительного обучения и затрат на подготовку тестовых данных. Соответственно, система, работающая в логическом подходе, сама может быть использована для автоматизации процесса подготовки обучающих данных. Степень автоматизации в данном случае будет пропорциональна показателям релевантности результатов автоматического анализа.

Таким образом, рассуждения на основе структурированных прецедентов позволяют уточнять результаты классификации и повышать достоверность распознавания объектов на космоснимках.

Благодарности

Работа выполнена за счёт гранта Российского научного фонда – РФФ (проект № 18-71-00109).

Литература

1. **Maggiori, E.** Can semantic labeling methods generalize to any city? The Inria Aerial Image Labeling Benchmark / E. Maggiori, Y. Tarabalka, G. Charpiat, P. Alliez // IEEE International Geoscience and Remote Sensing Symposium (IGARSS). – 2017. – P. 3226-3229.
2. **Hamaguchi, R.** Building detection from satellite imagery using ensemble of size-specific detectors / R. Hamaguchi, S. Hikosaka // The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops. – 2018. – P. 187-191.
3. **Zhang, A.** Building detection from satellite images on a global scale [Electronical Resource] / A. Zhang, X. Liu, A. Gros, T. Tiede // 30th Conference on Neural Information Processing Systems (NIPS 2016). – 2017. – arXiv preprint. – URL: <https://arxiv.org/abs/1707.08952> (request date 29.04.2020).
4. **Badrinarayanan, V.** SegNet: A deep convolutional encoder-decoder architecture for image segmentation / V. Badrinarayanan, A. Kendall // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2017. – Vol. 39, Issue 12. – P. 2481-2495.
5. **Filin, O.** Road detection with EOSResUNet and post vectorizing algorithm / O. Filin, A. Zapara, S. Panchenko // The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops. – 2018. – P. 211-215.
6. **Hamid, R.** Global-scale object detection using satellite imagery / R. Hamid, S. O'Hara, M. Tabb // The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. – 2014. – Vol. XL-3. – P. 107-113.
7. **Huang, X.** A multidirectional and multiscale morphological index for automatic building extraction from multispectral GeoEye-1 imagery / X. Huang, L.P. Zhang // Photogrammetric Engineering and Remote Sensing. – 2011. – Vol. 77, Issue 7. – P. 721-732.
8. **Zhang, Q.** A morphological building detection framework for high-resolution optical imagery over urban areas / Q. Zhang, X. Huang, G.X. Zhang // IEEE Geoscience and Remote Sensing Letters. – 2016. – Vol. 13. – P. 1388-1392.
9. **You, Y.** Building detection from VHR remote sensing imagery based on the morphological building index / Y. You, S. Wang, Y. Ma, G. Chen, B. Wang, M. Shen, W. Liu // Remote Sensing. – 2018. – Vol. 10, Issue 8. – 1288.
10. **Gurevich, I.B.** Descriptive Image Analysis: Genesis and Current Trends / I.B. Gurevich, V.V. Yashina // Pattern Recognition and Image Analysis. – 2017. – Vol. 27, Issue 4. – P. 653-674.
11. **Asatryan, D.G.** Gradient-based technique for image structural analysis and applications / D.G. Asatryan // Computer Optics. – 2019. – Vol. 43(2). – P. 245-250. – DOI: 10.18287/2412-6179-2019-43-2-245-250.
12. **Краснобаев, Е.А.** Сравнение бинарных дескрипторов особых точек изображений в условиях искажений / Е.А. Краснобаев, Д.В. Чистобаев, А.Л. Малышев // Компьютерная оптика. – 2019. – Т. 43, № 3. – С. 434-445. – DOI: 10.18287/2412-6179-2019-43-3-434-445.
13. **Фу, К.С.** Структурные методы в распознавании образов / К.С. Фу; пер. с англ. – М.: Мир, 1977. – 319 с.
14. **Kasimov, D.R.** Approximation of color images based on the clusterization of the color palette and smoothing boundaries by splines and arcs / D.R. Kasimov,

- A.V. Kuchuganov, V.N. Kuchuganov, P.P. Oskolkov // Programming and Computer Software. – 2018. – Vol. 44, Issue 5. – P. 295-302.
15. **Zadeh, L.A.** The concept of a linguistic variable and its application to approximate reasoning—I / L.A. Zadeh // Information Sciences. – 1975. – Vol. 8, Issue 3. – P. 199-249.
16. **Kuchuganov, A.V.** Recursions in image analysis problems / A.V. Kuchuganov // Pattern Recognition and Image Analysis. – Pleiades Publishing, Ltd., 2009. – Vol. 19(3). – P. 501-507.
17. **Baader, F.** Decidability and complexity of fuzzy description logics / F. Baader, S. Borgwardt, R. Peñaloza // Künstliche Intelligenz. – 2017. – Vol. 31(1). – P. 85-90.
18. **Yan, J.** A short survey of recent advances in graph matching / J. Yan, X.-C. Yin, W. Lin, C. Deng, H. Zha, X. Yang // Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval. – 2016. – P. 167-174.
19. **Rajput, M.K.** A survey on subgraph matching algorithm for graph database / M.K. Rajput, S. Kamalapur // International Journal for Scientific Research & Development. – 2016. – Vol. 3, Issue 12. – P. 149-152.
20. **Fernandez-Moral, E.** A new metric for evaluating semantic segmentation: leveraging global and contour accuracy / E. Fernandez-Moral, R. Martins, D. Wolf, P. Rives // Workshop on Planning, Perception and Navigation for Intelligent Vehicles, PPNIV17. – 2018. – P. 1051-1056.
21. **Blaschke, T.** Geographic object-based image analysis – Towards a new paradigm / T. Blaschke, G.J. Hay, M. Kelly, S. Lang, P. Hofmann, E. Addink, R.Q. Feitosa, F. Meer, H. Werff, F. Coillie, D. Tiede // ISPRS Journal of Photogrammetry and Remote Sensing. – 2014. – Vol. 87. – P. 180-191.

Сведения об авторе

Касимов Денис Рашидович, 1989 года рождения, в 2012 году окончил Ижевский государственный технический университет имени М.Т. Калашникова (ИжГТУ имени М.Т. Калашникова) по направлению «Информатика и вычислительная техника», специализация «Информационно-управляющие системы». Кандидат технических наук (2017 г.), старший научный сотрудник научно-образовательного центра «Интеллектуальные компьютерные технологии», доцент кафедры автоматизированных систем обработки информации и управления ИжГТУ имени М.Т. Калашникова. Область научных интересов: распознавание образов, поиск по содержанию, моделирование рассуждений, базы знаний, структурный анализ изображений. E-mail: kasden@mail.ru.

ГРНТИ: 28.23.15

Поступила в редакцию 11 марта 2020 г. Окончательный вариант – 7 мая 2020 г.

Searching and describing objects in satellite images on the basis of modeling reasoning

D.R. Kasimov¹

¹*Kalashnikov Izhevsk State Technical University,
426069, Izhevsk, Russia, Studencheskaya 7*

Abstract

The article presents an approach to a problem of contextual search and description of objects in raster satellite images, which consists in modeling reasoning on the basis of structured cases. As a result of image processing, an adjacency graph of color regions is constructed. The object is characterized by color, attributes of the form of segments of the border and the shape of the object as a whole. A structured case is represented in the form of a beam graph, whose arcs are ordered according to a positive bypass of the region boundaries. Using a graph matching algorithm, occurrences of cases stored in the system database are detected in the analyzed image. When the occurrence is detected, a case-based inference rule is applied. The degree to which an object belongs to a certain class depends not only on the properties of the object itself, but also on the reliability of the surrounding objects. The contextual search strategy contains stages of recursion and iteration. In contrast to neural network technologies, the proposed approach allows one not only to classify image objects, but also to form their structured descriptions. In addition, the classification decision issued by the system has a reasoned justification. The results of the experiment show that reasoning based on structured cases allows refining the results of classification and increasing the reliability of object recognition in satellite images.

Keywords: computer vision, digital image processing, pattern recognition, structural analysis, segmentation, approximation, adjacency graph, beam graph, case-based reasoning.

Citation: Kasimov DR. Searching and describing objects in satellite images on the base of modeling reasoning. *Computer Optics* 2020; 44(5): 772-781. DOI: 10.18287/2412-6179-CO-716.

Acknowledgements: The research was financially supported by the Russian Science Foundation (Project No. 18-71-00109).

References

- [1] Maggiori E, Tarabalka Y, Charpiat G, Alliez P. Can semantic labeling methods generalize to any city? The Inria Aerial Image Labeling Benchmark. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS) 2017*; 3226-3229.
 - [2] Hamaguchi R, Hikosaka S. Building Detection From Satellite Imagery Using Ensemble of Size-Specific Detectors. *IEEE Conf on Computer Vision and Pattern Recognition (CVPR) Workshops 2018*; 187-191.
 - [3] Zhang A, Liu X, Gros A, Tiede T. Building Detection from Satellite Images on a Global Scale. *30th Conference on Neural Information Processing Systems (NIPS 2016) 2017*. Source: (<https://arxiv.org/abs/1707.08952>).
 - [4] Badrinarayanan V, Kendall A. SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans Pattern Anal Mach Intell* 2017; 39(12): 2481-2495.
 - [5] Filin O, Zapara A, Panchenko S. Road detection with EOSResUNet and post vectorizing algorithm. *IEEE Conf on Computer Vision and Pattern Recognition (CVPR) Workshops 2018*; 211-215.
 - [6] Hamid R, O'Hara S, Tabb M. Global-scale object detection using satellite imagery. *ISPRS Archives* 2014; XL-3: 107-113.
 - [7] Huang X, Zhang LP. A multidirectional and multiscale morphological index for automatic building extraction from multispectral GeoEye-1 imagery. *Photogramm Eng Remote Sensing* 2011; 77(7): 721-732.
 - [8] Zhang Q, Huang X, Zhang GX. A morphological building detection framework for high-resolution optical imagery over urban areas. *IEEE Geosci Remote Sens Lett* 2016; 13: 1388-1392.
 - [9] You Y, Wang S, Ma Y, Chen G, Wang B, Shen M, Liu W. Building detection from VHR remote sensing imagery based on the morphological building index. *Remote Sens* 2018; 10(8), 1288.
 - [10] Gurevich IB, Yashina VV. Descriptive Image Analysis: Genesis and Current Trends. *Pattern Recognition and Image Analysis* 2017; 27(4): 653-674.
 - [11] Asatryan DG. Gradient-based technique for image structural analysis and applications. *Computer Optics* 2019; 43(2): 245-250. DOI: 10.18287/2412-6179-2019-43-2-245-250.
 - [12] Krasnabayeu YA, Chistabayeu DV, Malyshev AL. Comparison of binary feature points descriptors of images under distortion conditions. *Computer Optics* 2019; 43(3): 434-445. 10.18287/2412-6179-2019-43-3-434-445.
 - [13] Fu KS. *Syntactic methods in pattern recognition*. New York, London: Academic Press; 1974.
 - [14] Kasimov DR, Kuchuganov AV, Kuchuganov VN, Oskolkov PP. Approximation of color images based on the clusterization of the color palette and smoothing boundaries by splines and arcs. *Program. Comput Softw* 2018; 44(5): 295-302.
 - [15] Zadeh LA. The concept of a linguistic variable and its application to approximate reasoning—I. *Inf Sci* 1975; 8(3): 199-249.
 - [16] Kuchuganov AV. Recursions in image analysis problems. *Pattern Recognit Image Anal* 2009; 19(3): 501-507.
 - [17] Baader F, Borgwardt S, Peñaloza R. Decidability and complexity of fuzzy description logics. *Künstliche Intelligenz* 2017; 31(1): 85-90.
-

-
- [18] Yan J, Yin XC, Lin W, Deng C, Zha H, Yang X. A short survey of recent advances in graph matching. Proc 2016 ACM Int Conf on Multimedia Retrieval 2016; 167-174.
- [19] Rajput MK, Kamalapur S. A survey on subgraph matching algorithm for graph database. International Journal for Scientific Research & Development 2016; 3(12): 149-152.
- [20] Fernandez-Moral E, Martins R, Wolf D, Rives P. A new metric for evaluating semantic segmentation: leveraging global and contour accuracy. Workshop on Planning, Perception and Navigation for Intelligent Vehicles (PPNIV17) 2018. 1051-1056
- [21] Blaschke T, Hay GJ, Kelly M, Lang S, Hofmann P, Addink E, Feitosa RQ, Meer F, Werff H, Coillie F, Tiede D. Geographic object-based image analysis – Towards a new paradigm. ISPRS J Photogramm Remote Sens 2014; 87: 180-191.
-

Author's information

Denis Rashidovich Kasimov (b. 1989) graduated from Kalashnikov Izhevsk State Technical University (ISTU) in 2012, majoring in Computer Science and Engineering, Information and Control Systems. He received his PhD in 2017. Currently he works as the senior researcher at the research and education center "Intelligent Computer Technologies" and as an associate professor at Automated Data Processing and Control Systems department of ISTU. Research interests are pattern recognition, content-based search, modeling reasoning, knowledge base, and structural image analysis. E-mail: kasden@mail.ru.

Received March 11, 2020. The final version – May 6, 2020.
