

Обнаружение строений по данным мультиспектральной космической съемки и цифровой модели поверхности

А.В. Дунаева^{1,2}, Ф.А. Корнилов¹

¹Институт математики и механики им. Н.Н. Красовского УрО РАН, С. Ковалевской 16, Екатеринбург, Россия, 620990

²Уральский федеральный университет им. первого Президента России Б.Н. Ельцина, Мира 19, Екатеринбург, Россия, 620002

Аннотация. Работа посвящена обнаружению строений на спутниковых снимках земной поверхности с использованием данных о высотах рельефа и объектов местности. Рассматриваются два различных подхода к решению данной задачи. Первый основан на выделении высотных объектов местности по карте высот и уточнении их границ на основе сегментации спектральной информации. Второй подход представляет собой использование сверточной нейронной сети специального вида. Приводятся результаты сравнения качества работы рассмотренных подходов на реальных данных.

1. Введение

В настоящее время мониторинг состояния земной поверхности на основе аэрокосмических изображений играет важную роль во многих прикладных областях: сельском хозяйстве, строительстве, анализе чрезвычайных ситуаций и других. Такой подход обладает рядом неоспоримых преимуществ: своевременность, широта охвата территорий и относительно невысокая стоимость. В то же время, многие задачи решаются в ручном режиме оператором, что делает актуальной разработку средств автоматического извлечения из снимков требуемой информации.

Дешифрирование и классификация объектов местности являются одними из важнейших задач при обработке спутниковых снимков земной поверхности. Решение подобных задач находит свое применение в картографии (обновление топографических карт), обнаружении изменений в составе объектов местности (анализ городской застройки, незаконные вырубки леса), навигации летательных аппаратов по геофизическим полям и других. Однако существующие на данный момент методы не позволяют классифицировать объекты местности по космическим снимкам с достаточной точностью. Это связано, в первую очередь, со сложностью и вариативностью рассматриваемых сцен, содержащей огромное количество объектов различной природы, подверженных существенному влиянию условий съемки. Все это обуславливает значимость исследований, направленных на преодоление указанных трудностей и разработку проблемно ориентированных программных комплексов обработки данных дистанционного зондирования.

В данной работе для повышения качества классификации предполагается использовать цифровую модель поверхности, содержащую данные о высотах рельефа и объектов местности. Эту информацию можно получить путем съемки сцены с разных ракурсов. Два таких снимка

образуют стереопару, передаваемую на вход алгоритму стереосопоставления, который вычисляет матрицу сдвигов пикселей между кадрами, принадлежащих одному и тому же объекту. Далее на основе матрицы сдвигов и данных о параметрах съемки строится цифровая модель поверхности путем вычисления высоты объектов. В данной работе использовалась цифровая модель поверхности, полученная в результате работы алгоритма, предложенного в [1]. Поскольку основной интерес представляла городская застройка, то для стереосопоставления были использованы панхроматические снимки с разрешением 0.5 м, а разрешение итоговой модели составило 1 м на пиксель.

Работа организована следующим образом. Во втором разделе описан подход к задаче обнаружения строений, основанный на выделении высотных объектов местности по карте высот и уточнении их границ на основе сегментации спектральной информации. Третий раздел посвящен сверточной нейронной сети специальной архитектуры, предназначенной для выделения на изображениях объектов заданного класса. В четвертом разделе приводятся результаты сравнения рассмотренных подходов на реальных данных.

2. Сегментация изображений и выделение высотных объектов

Сегментация применяется для предварительной обработки изображений во многих задачах компьютерного зрения, так как позволяет проводить анализ не отдельных пикселей, а однородных и целостных по некоторому признаку областей. Для разбиения изображения, могут использоваться различные признаки, такие как цвет, текстура и т.д. К настоящему времени разработано большое количество алгоритмов сегментации. В работах [2,3] приведено описание и сравнение некоторых из них применительно к задаче поиска строений. Рассмотренные алгоритмы сегментации демонстрируют высокое качество работы, но не используют информацию о высотах объектов.

Для оценки влияния данных о высоте местности на результат сегментации спутниковых снимков в [3] был представлен алгоритм, основанный на идее объединения пикселей в области на основе близости их яркостей и высот. Приведем его краткое описание. На вход алгоритм получает четырехканальный снимок и карту высот. Выполняется следующая последовательность шагов.

- Медианная фильтрация спектрального изображения.
- Предварительное выделение области. Выполняется перебор всех пикселей изображения. Начиная с рассматриваемого пикселя выполняется рекурсивный поиск соседних пикселей, имеющих близкую высоту и яркость (в четырехмерном пространстве). Далее по найденной области вычисляются значения мод высоты и каждого из четырех каналов по отдельности.
- Выделение области: в полученной области выбирается точка, имеющая высоту, в точности равную вычисленной моде высоты, а яркость по каждому каналу – наиболее близкую к модам по этим каналам. Такое решение обусловлено тем, что моды по спектральным и высотным данным могут не совпадать. Далее от найденной точки строим новую область, аналогично п.2, и закрашиваем её цветом вычисленной моды
- Повтор шагов 2-3. Из-за зашумленности входных данных области, полученные данным алгоритмом, будут иметь небольшую площадь. Итеративное повторение процедуры сегментации позволит объединить области в более крупные.
- Фильтрация областей малой площади. Из-за выбросов яркости или поля высот могут оставаться области, имеющие пренебрежимо малую площадь. Такие области объединяются с соседними таким образом, чтобы область, к которой добавляются точки, имела большую площадь, а её высота и яркость была настолько близка к добавляемым точкам, насколько это возможно. Данная процедура выполняется для областей в порядке возрастания их площади.

Выполнив сегментацию снимка, необходимо определить принадлежность полученных областей к зданиям и сооружениям. Для этого необходимо по карте высот выделить так

называемые высотные объекты – т.е. те объекты, которые возвышаются над рельефом. К таковым относятся строения и лесные массивы. Для этого предлагается следующий алгоритм.

1. *Поиск перепадов высот.* Для каждого пикселя p карты высот со значением высоты h рассматриваются его одномерные окрестности $O_d^{\rightarrow}(p)$ и $O_d^{\downarrow}(p)$ радиуса d в горизонтальном и вертикальном направлении. Для пикселей $p_i \in O_d(p)$ вычисляются $\Delta h_i = h_i - h_{i+1}, i \in [-d, d - 1]$, где индекс 0 соответствует центру окрестности – точке p . И если $\Delta h_0 = \max_i \Delta h_i$ или $\Delta h_0 = \min_i \Delta h_i$, и $|\Delta h_0| \geq T$, то точка p отмечается как имеющая значимый перепад высоты. Точки, на которых одновременно достигается минимум и максимум не рассматриваются, так как являются выбросами на карте высот. Параметр T отвечает за настройку алгоритма для обнаружения строений требуемой высоты: например, с его помощью можно исключать невысокие строения, относящиеся к частному сектору.

2. *Выделение высотных областей.* По точкам, имеющим значимый перепад, выполняется построение отрезков, точки которых и будут являться найденными высотными областями. Для построения отрезков на карте высот по горизонтали и вертикали в прямом и обратном направлении производится поиск точек с перепадом высоты вверх, после чего от найденной точки проводится отрезок, который идет:

- до точки с перепадом высоты вверх, значение которого добавляется к значению начального перепада, после чего отрезок продолжается;
- до точки, в которой значение высоты меньше или близко к значению начала отрезка;
- до накопления максимально возможной длины отрезка.

Выполнение данной процедуры в прямом и обратном направлении (слева направо и наоборот) позволяет выделить даже те области объектов на карте высот, у которых из-за воздействия шума может быть размыта одна из сторон.

Далее отбрасываются точки, найденные только по горизонтали или только по вертикали. Оставшиеся точки образуют связанные области, среди которых выбираются лишь те, что содержат хотя бы одну точку со значимым перепадом высот. Такая операция позволяет отсеять ложные пересечения вертикальных и горизонтальных отрезков.

3. *Уточнение формы найденных областей.* На карте высот рассматриваются односвязные области, все точки которых имеют одинаковую высоту. Если больше половины точек области были отмечены как высотные на предыдущем шаге, то и оставшиеся точки также отмечаются как высотные. Таким образом выравнивается граница найденных областей и отбрасываются ложные срабатывания, связанные со склонами холмов и иными резкими перепадами рельефа местности. Области малого размера удаляются из рассмотрения.

4. *Выделение зданий среди найденных высотных объектов.* К обнаруженным высотным объектам относятся не только здания и сооружения, но и высотная растительность (деревья). Эффективно отделить растительность от остальных объектов на спутниковых снимках позволяет индекс NDVI [4], который вычисляется по красному и инфракрасному каналам спутникового снимка.

Таким образом, получается двухэтапный алгоритм. Выполнение сегментации позволяет разбить снимок на однородные области, однако, при этом здания могут быть покрыты несколькими областями. Выделение высотных объектов позволяет найти объект целиком, но с невысокой точностью локализации границ. Предлагается объединить результаты алгоритмов поиска высотных объектов и сегментации следующим образом: если более чем 75% пикселей сегментированной области отмечены как высотные, то и всю область считать высотной. Это позволяет повысить точность локализации объектов и уменьшить количество ложных тревог (на участках деревьев, близко примыкающих к зданиям). Результат работы предложенного подхода представлен на рис. 1 и в таблице.

3. Применение сверточных нейронных сетей для обнаружения строений на спутниковых снимках

Использование сверточных нейронных сетей позволило значительно повысить качество работы алгоритмов компьютерного зрения при решении самых разных задач. Не стала исключением и

задача обнаружения и классификации объектов на спутниковых снимках. В связи с большим интересом различных организаций к решению этой задачи, за последний год было проведено как минимум четыре крупных онлайн-соревнования [5-8] по классификации объектов на снимках земной поверхности. Среди многообразия решений лучшее качество работы для обнаружения строений в [5,6] показали сверточные нейронные сети - модификации сети U-Net [9]. Стоит отметить, что наибольшая точность классификации объектов достигается путем обучения отдельной модели сети U-Net для каждого класса объектов. Далее будет рассмотрена модификация, предложенная в [10], а также влияние на качество ее работы дополнительной информации в виде карты высот.

U-Net принимает на вход многомерный массив, составленный из каналов спутникового снимка, индекса растительности CCCI и водного индекса NDWI, вычисленных по каналам снимка. В виду ограниченности памяти GPU, снимок обрабатывается фрагментами 112×112 пикселей. Обработка данных в U-Net выполняется двумя частями сети: первая предназначена для извлечения признаков и определения класса объекта (энкодер), вторая (декодер) для получения итоговой разметки снимка на два класса – здания и фон.

Для оценки качества работы нейронной сети в данной работе было выполнено ее обучение на вручную размеченном четырехканальном спутниковом снимке участка земной поверхности размером порядка 5 × 5 км, и содержащем более 3000 строений – многоэтажные постройки, гаражи, ангары и частные дома. Так как работа посвящена поиску объектов одного класса, не использовался водный индекс NDWI, а вместо индекса растительности CCCI был вычислен индекс растительности NDVI. Для оценки степени влияния данных на качество распознавания, было обучено четыре модели U-Net на следующих наборах данных:

- четыре канала спутникового снимка, индекс NDVI, карта высот;
- четыре канала, карта высот;
- четыре канала, индекс NDVI;
- четыре канала.

Результат работы обученных сетей представлен на рис. 1 и в таблице.

4. Результаты работы и сравнение предложенных подходов

Оценка качества работы алгоритмов проводилась на фрагменте четырехканального спутникового снимка размером 900×500 пикселей с пространственным разрешением 1 метр на пиксель. Области, отмеченные алгоритмами как здания, сравнивались с ручной разметкой по следующим критериям:

1. Мера пересечения на объединение (*Intersection over Union*), учитывающая точность локализации обнаруженных объектов

$$IoU = \frac{|S_{obj} \cap R|}{|S_{obj} \cup R|},$$

где S_{obj} – найденные области строений, а R – ручная разметка.

2. Процент обнаруженных строений N_{obj} .

3. Точность = $\frac{|S_{obj} \cap R|}{|S_{obj}|}$, полнота = $\frac{|S_{obj} \cap R|}{|R|}$ и ложное обнаружение = $\frac{|S_{obj} \setminus R|}{|S \setminus R|}$, где S – весь снимок.

Полученные результаты представлены в таблице и на рисунке 1.

Из двух рассмотренных подходов обнаружения строений лучшее качество работы продемонстрировала сеть U-Net. Ее преимуществом по сравнению с алгоритмом, выделяющим объекты по карте высот, является способность выделять частные дома и другие малоэтажные постройки, незаметные на карте высот – а это, как минимум, треть объектов на исходном снимке 5 × 5 км. Также нейронные сети позволяют более точно выделять границы строений.

Таблица 1. Сравнение качества работы алгоритмов обнаружения строений.

Алгоритм	IoU	N_{obj}	Точность попиксельно	Полнота попиксельно	Ложное обнаружение
Высотные объекты + сегментация	0.56	0.86	0.68	0.76	0.085
U-Net (4 канала, ndvi, высота)	0.72	0.90	0.79	0.90	0.062
U-Net (4канала, высота)	0.74	0.85	0.82	0.88	0.050
U-Net (4канала, ndvi)	0.72	0.80	0.80	0.88	0.059
U-Net (4канала)	0.74	0.82	0.86	0.85	0.037

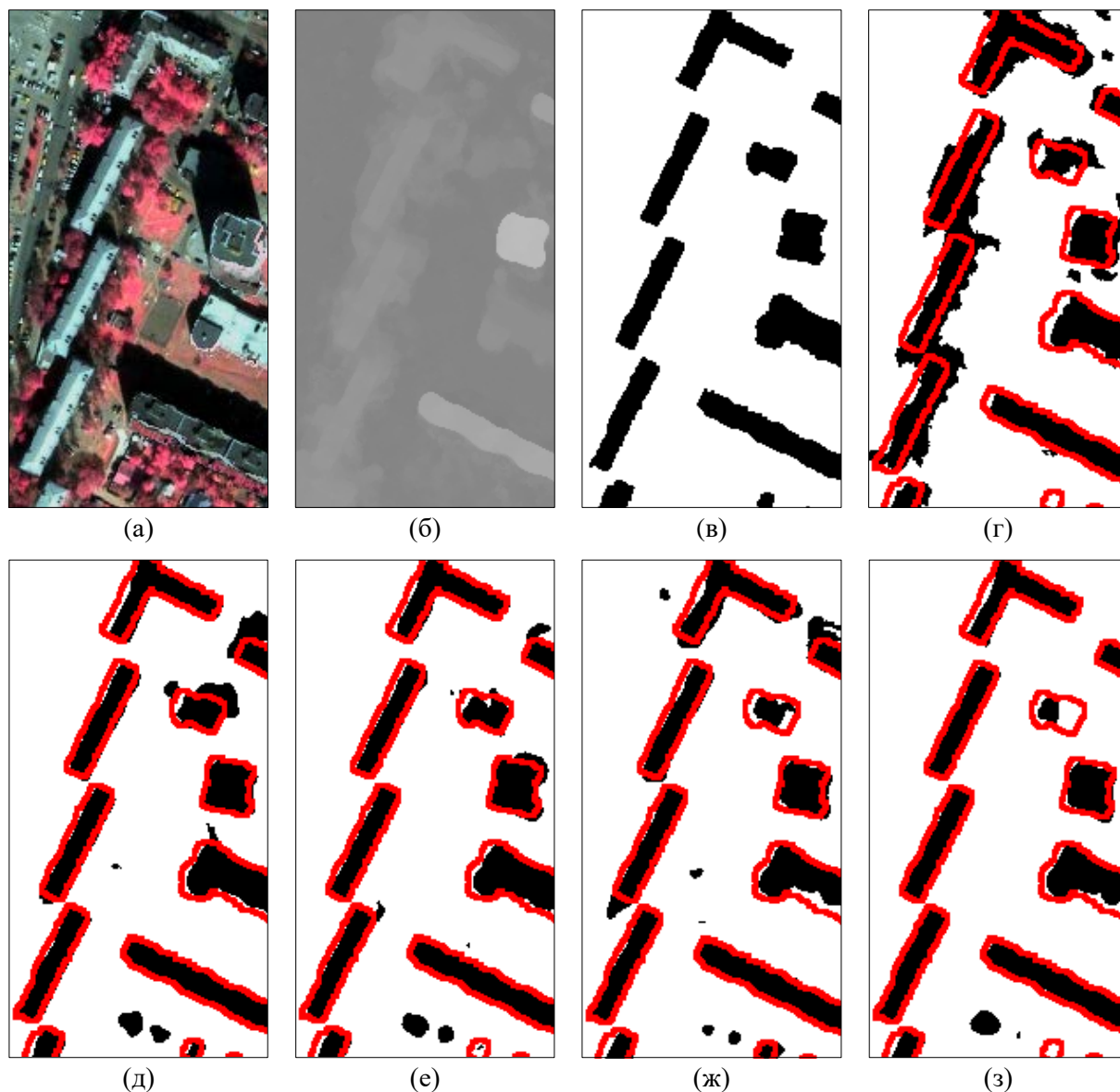


Рисунок 1. (а) Исходное изображение (ближний ИК, красный и зеленый каналы), (б) карта высот, (в) ручная разметка строений, (г) поиск высотных объектов с уточнением границы строений посредством сегментации. Обнаружение строений сетью U-Net с обучением: (д) по четырем каналам, матрице высот и индексу NDVI, (е) по четырем каналам и матрице высот, (ж) по четырем каналам и индексу NDVI, (з) с обучением только по четырем каналам снимка.

Из полученных данных можно видеть, что использование информации о высоте, позволяет U-Net обнаруживать большее количество зданий. Однако, сети, обучавшиеся без высотной

информации, качественнее выполняют выделение границ объектов, о чем свидетельствует более высокое значение IoU . Предположительно это связано с тем, что при наличии высотной информации сеть опирается на нее для выделения границ строений. Вероятно, повысить точность распознавания границ при использовании высотной информации можно за счет изменения архитектуры сети или других параметров ее обучения. Использование индекса NDVI привело к неоднозначным результатам: совместное использование NDVI и карты высот позволило повысить качество распознавания, в то время как обучение только на четырех каналах снимка привело к лучшим результатам, чем при обучении на снимке и NDVI.

5. Заключение

В данной работе были рассмотрены два подхода к обнаружению и локализации строений на мультиспектральных космических снимках земной поверхности с использованием карты высот. Наилучшее качество работы показал подход, использующий нейронную сеть U-Net, однако его применение возможно только при наличии достаточно большого набора данных для обучения сети. В дальнейшем предполагается рассмотреть другие архитектуры нейронных сетей, а также исследовать влияние размера и состава обучающего набора на качество обнаружения строений. Для уточнения границы найденных объектов предполагается использовать контурный анализ аналогично методу, предложенному в [11].

6. Благодарности

Работа выполнена при финансовой поддержке комплексной программы ФНИ УрО РАН "Классические и современные методы теории аппроксимации и приложения" №18-1-1-14.

7. Литература

- [1] Костоусов, В.Б. Создание цифровой модели местности по данным космической стереосъемки / В.Б. Костоусов, Д.С. Перевалов, Ф.А. Корнилов // Материалы XXX конференции памяти выдающегося конструктора гироскопических приборов Н.Н. Острякова. – СПб.: АО "Концерн "ЦНИИ Электроприбор", 2016. – С. 382-388.
- [2] Марьянова, А.В. Исследование качества алгоритмов сегментации в зависимости от размера объектов на изображении / А.В. Марьянова // Труды 46-й Международной молодежной школы-конференции "Современные проблемы математики и её приложений". – Екатеринбург, 2015. – С.129-134.
- [3] Дунаева, А.В. Обнаружение строений на спутниковых снимках с использованием карты высот / А.В. Дунаева, Ф.А. Корнилов // Современные проблемы математического моделирования, обработки изображений и параллельных вычислений 2017: труды Междунар. науч. конф. Том II. Донской гос. техн. ун-т. – Ростов-на-Дону: ООО «ДГТУ-Принт», 2017. – С. 69-77.
- [4] Козодеров, В.В. Методы оценки состояния почвенно-растительного покрова по данным оптических систем дистанционного аэрокосмического зондирования / В.В. Козодеров, Т.В. Кондранин. - Москва, МФТИ, 2008. – С. 222.
- [5] Dstl Satellite Imagery Feature Detection Competition Web Page [Electronic resource]. - Access mode: <https://www.kaggle.com/c/dstl-satellite-imagery-feature-detection> (01.12.2017).
- [6] The SpaceNet Challenge Web Page [Electronic resource]. - Access mode: <https://crowdsourcing.topcoder.com/spacenet> (01.12.2017).
- [7] Use Satellite Data to Track the Human Footprint in the Amazon Rainforest [Electronic resource]. - Access mode: <https://www.kaggle.com/c/planet-understanding-the-amazon-from-space> (01.12.2017).
- [8] Safe Passage: Detecting and Classifying Vehicles in Aerial Imagery rainforest [Electronic resource]. – Access mode: <https://www.datasciencechallenge.org/challenges/1/safe-passage> (01.12.2017).
- [9] Ronneberger O. U-Net: Convolutional networks for biomedical image segmentation / O. Ronneberger, P. Fischer, T. Brox // International Conference on Medical Image Computing and

- Computer-Assisted Intervention. MICCAI 2015. Lecture Notes in Computer Science. Springer International Publishing. – 2015. – Vol 9351. – P. 324-241.
- [10] Iglovikov, V. Satellite Imagery Feature Detection using Deep Convolutional Neural Network: A Kaggle Competition / V. Iglovikov, S. Mushinskiy, V. Osin // June 2017. [Electronic resource]. – Access mode: <https://arxiv.org/abs/1706.06169> (01.12.2017).
- [11] Partovi T. Statistical Buildings Roof Reconstruction from WorldView-2 Stereo Imagery / T. Partovi, H. Huang, T. Kraub, H. Mayer, P. Reinartz // The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. – 2015 – Vol. XL-3/W2. – P. 161-167.

Building detection from satellite multispectral images using a digital surface model

A.V. Dunaeva^{1,2}, F.A. Kornilov¹

¹N.N. Krasovskii Institute of Mathematics and Mechanics of the Ural Branch of the Russian Academy of Sciences, S. Kovalevskaya street 16, Yekaterinburg, Russia, 620990

²Ural Federal University named after the first President of Russia B.N. Yeltsin, Mira street 19, Yekaterinburg, Russia, 620002

Abstract. We consider two approaches to building detection from the satellite imagery using height data of objects and relief obtained from the digital surface model. The first is based on the selection of high-rise objects on a terrain height map and then improves the recognition accuracy using the segmentation of spectral information. The second approach uses a convolutional neural network of a special kind. The results of the quality comparison of the algorithms on real data are presented.

Keywords: Image processing, remote sensing imagery, digital surface model, building detection, neural networks.