

УДК 528.873:004.932.2

АЛГОРИТМ МОРФОЛОГИЧЕСКОГО АНАЛИЗА И ОПРЕДЕЛЕНИЯ ПОРЯДКОВ ВОДОТОКОВ НА ОСНОВЕ ЦИФРОВОЙ МОДЕЛИ РЕЛЬЕФА ДЛЯ РАЗРАБОТКИ ОТКРЫТЫХ ГЕОИНФОРМАЦИОННЫХ СИСТЕМ

^{1,2}Шевырёв С.Л.¹Дальневосточный геологический институт ДВО РАН, Владивосток, e-mail: shevirev@mail.ru;²Дальневосточный федеральный университет, Владивосток

Наличие свободно распространяемой спутниковой цифровой модели рельефа (ЦМР) SRTM на большую часть территории земного шара делает этот источник дистанционных данных привлекательным для ведения широкого спектра прикладных и исследовательских работ. Методы автоматизированного распознавания и установления иерархии элементов эрозионной сети на основе SRTM применяются в географических информационных системах для оценки степени эрозионного расчленения территории, при неотектонических и ландшафтных реконструкциях, а также для прогноза минерально-сырьевых ресурсов и различного картоирования. Несмотря на активные темпы развития информационных технологий и методов анализа изображений в настоящее время, выделение гидросети и разбивка ее на генерации (порядки) водотоков все еще представляют собой нетривиальную задачу, решение которой осложнено недостатками цифровой основы и несовершенством применяемых алгоритмов распознавания или ограничениями их применения. Погрешности в определении абсолютных отметок и оставление на усмотрение пользователя ответственности за принимаемые решения также снижают достоверность результата. В настоящей работе на примере инструментария языка Python рассматривается предлагаемый автором простой и эффективный алгоритм дешифрирования гидросети и подсчет числа Стралера (порядка) водотока с использованием морфологической обработки, реализуемой в открытых системах научных расчетов без устанавливаемых пользователем пороговых значений. Формат вывода результата, включающий пространственную привязку, может быть использован для его компоновки в географических информационных системах (QGIS, ArcGIS и других).

Ключевые слова: морфологический анализ изображений, геоинформатика, цифровая модель рельефа, дистанционное зондирование Земли, число Стралера

ALGORITHM OF MORPHOLOGICAL ANALYSIS AND IDENTIFICATION OF STREAM ORDER IN DIGITAL ELEVATION MODEL FOR DEVELOPMENT OF OPEN-SOURCE GEOINFORMATIC SYSTEMS

^{1,2}Shevirev S.L.¹Far East Geological Institute, Vladivostok, e-mail: shevirev@mail.ru;²Far Eastern Federal University, Vladivostok

Nowadays, the freely available satellite digital relief model (DRM) for the most part of the globe makes this source of remote sensing data affordable for the great variety of applied activities and the research. Methods of remote detecting of erosion network elements and establishing of their hierarchy are applicable for assessment of the erosion dissection grade, neotectonic and landscape reconstructions as well as for resource management and various mapping. Despite of rapid development of information technologies and image analysis recently, extraction of erosion network and dividing it into the generations (orders) remains challenging task, which is complicated by artifacts in digital relief model and shortcomings in current detection algorithms or by their limitations. Errors in absolute elevations and leaving decision making onto the user responsibility also decreases reliability of available methods. Given paper considers simple and efficient algorithm of erosion network detecting and computing of Strahler number with morphological threshold-less processing in open-source scientific computing environments on a sample of Python. Output result containing spatial reference could be used for its stacking in geographic information systems (QGIS, ArcGIS etc.).

Keywords: morphological image processing, geoinformatics, digital relief model, remote sensing of the Earth, Strahler number

Оперативность в принятии природо-пользовательских и природоохранных решений при освоении и эксплуатации территорий требует наращивания имеющегося вычислительного и методического аппарата. Данные цифровой радарной интерферометрической модели рельефа SRTM, полученной совместными усилиями NASA и ESA, представляют собой массив высот в метрах с пространственным разрешением до 90 м, распространяемый в виде фрагмен-

тов мозаики шириной 5×5 градусов и доступный на всю территорию планеты до 60 С.Ш. и до 54 Ю.Ш. [1, 2]. Компьютерная обработка и анализ SRTM является источником разнородной информации о территории покрытия, которая может быть извлечена из параметров рельефа, в том числе: устойчивость подстилающих пород к размыву, наличие неоднородностей и разрывов в массиве грунта, гидрологический режим рек и другие сведения.

Несмотря на доступность данных ЦМР широкому кругу исследователей, анализ эрозионной сети и определение порядков ее притоков по Стралеру [3], представляет собой нетривиальную задачу, успешное выполнение которой осложняется наличием артефактов SRTM и тем, что помимо рельефа, в качестве поверхности, отражающей радиоволны, могут выступать растительный слой и местные предметы. Целью настоящей работы является разработка альтернативного алгоритма дешифрирования эрозионной сети на основе SRTM с применением вычислительных средств с открытым исходным кодом. Для достижения поставленной цели необходимо решение задач: формирования потока обработки исходных данных; подбор необходимого математического аппарата для выполнения каждого этапа анализа, а также пробная имплементация алгоритма с разработкой программы-прототипа, выполняющей дешифрирование эрозионной сети по спутниковой ЦМР и ее экспорт в ГИС-совместимый формат.

Материалы и методы исследования

Для корректной постановки задач исследования необходимо принимать во внимание, что при описании эрозионной сети, выделяемой на спутниковой основе, речь идет не о буквальных «водотоках», а об их морфологическом выражении в цифровой модели рельефа – тальвегах, т.е. линиях, соединяющих наиболее пониженные участки эрозионной ложбины, вне зависимости от наличия или отсутствия в ней постоянного или временного потока воды.

В качестве существующих методов детектирования объектов эрозионной сети следует отметить: ручное дешифрирование; автоматизированное распознавание с помощью модуля Spatial Analysis ArcGIS [4–6], использование малоизвестного в отечественных исследованиях набора инструментов TopoToolbox проприетарной Mathworks Matlab [7], а также недавнюю разработку исследователей из Флоридского университета, использующую библиотеки Python Numpy (бесплатная программа, поддерживается сообществом) и ArcPy (лицензия ESRI) [8]. Основным недостатком перечисленных инструментов является их привязанность к проприетарным API Matlab и ArcGIS, ограничивающими их применением конечными пользователями. В связи с этим разрабатываемый алгоритм должен адекватно дешифрировать гидросеть с учетом вышеуказанных особенностей исход-

ных данных, разбивать детектированную сеть на притоки и определять их степень ветвления; желательно также не привязывать его имплементацию к коммерческой среде выполнения.

Сведения о распределении высот, представленные в модели SRTM, позволяют судить о морфологии отражающей поверхности и характеристиках подстилающего субстрата, обусловивших наблюдаемую картину. Эрозионное расчленение рельефа, его густота, степень и характер разветвленности свидетельствуют о постоянной или сезонной обводненности территории, а также о ее дренировании. Число Стралера позволяет установить иерархическую соподчиненность водотоков и порядок речных [3]. Этот параметр рассчитывается следующим образом: водотоки, не имеющие притоков, имеют первый порядок, те из них, в которые впадают исключительно водотоки первого порядка, имеют второй порядок, те, в которые впадает водоток второго порядка, имеют третий порядок и т.д.

Разрабатываемый алгоритм должен обладать необходимыми свойствами, включающими: имплементацию методик, описанных в открытых источниках, а также возможность выполнения проектируемых с ним приложений в свободных средствах программирования, средах и языках, что обеспечит повторяемость результатов. Адаптируемость алгоритма должна позволить применять его для анализа территорий, характеризующихся различными особенностями распределения высот. Кроме того, значения ошибок определения высот изображения, являющиеся особенностью SRTM [9], не должны оказывать существенного влияния на детектирование водотоков.

Существующие и описанные в литературе методы распознавания водотоков в самом общем случае сводятся к определению источников («точек начала течения»), мест впадения («точек слияния линейных водотоков друг с другом или с площадным водным объектом»). Применяется нахождение производных и градиентов поверхности рельефа, при этом водоток детектируется как область смены направления градиента поверхности рельефа φ:

$$grad \varphi = i \frac{\partial \varphi}{\partial x} + j \frac{\partial \varphi}{\partial y}$$

где *i, j* – масштабные коэффициенты.

Очевидная логичность данного метода при его практической реализации сталкивает-

ся с затруднением выделения границы направления градиента рельефа в «пойме», т.е. уплощённом днище ложбины эрозионного стока в связи с незначительным относительным перепадом высот, а также наличием помех для отражения: крон деревьев, кустарников и иных объектов. Этот факт снижает эффективность методов, основанных на выявлении направлений изменения высот модели, а также выявления границ градиента, а также делает привлекательными методы, основанные на анализе геометрии и морфологии радарной интерферометрической поверхности.

Предлагаемый алгоритм, использующий методы морфологического анализа, показан на схеме (рис. 1). Он включает два цикла, условно названных «последовательное затопление» и «поиск соседей».

В природе затопление речной долины приводит к формированию на ее участке водной поверхности, границы распространения которой повторяют очертания изолиний рельефа. При этом осевая линия этой поверхности, равноудалённая от изолиний, в начале затопления совпадает с руслом реки, а затем, по мере сокращения долины под водой, смещается к оси речной долины. При затоплении долин ее притоков схожая картина наблюдается и в них. При этом происходит естественное сглаживание небольших неровностей рельефа. Предлагаемая методология использует анализ эрозионной сети, выбор подмножества диапазона высот радарного изображения, морфологические операции утончения, наращивания и скелетизации. Указанные функции входят в библиотеку Scikit-image, предназначенную для анализа изображений в Python [10]. Выбор и маркирование отдельных водотоков осуществлялись с помощью операции `binary labels`, позволяющей присваивать уникальный идентификатор кластеру бинарного (черно-белого) изображения.

Тестовая имплементация алгоритма осуществлялась в открытой среде научного программирования Spyder на языке Python3 с использованием, помимо функций ядра Python, а также упоминавшейся выше библиотеки Scikit-image, библиотеки Numpy (численные расчёты) и gdal (Geographic data abstraction library), которые обеспечивают ввод, обработку и вывод численных данных с пространственной привязкой.

Алгоритм включает 4 функциональных блока, работа которых не требует указания пользователем каких-либо параметров, влияющих на результативность анализа, за исключением указания пути к файлу ис-

следуемого изображения (блок 1). Действия блока 2 выполняются циклически с шагом 1/10 перепада высот до достижения максимальной отметки и включают: «затопление», т.е. выбор и создание булевой матрицы для подмножества высот ЦМР, (2.1), морфологическое истончение *skimage.thin* (2.2), удаление части структурного рисунка, входящего в пределы предыдущих полей «затопления» (2.3), сохранение остаточного рисунка в матрице *нераспределенных водотоков* (2.4). По достижении максимальной высотной отметки осуществляется переход к блоку 3, минуя коррекцию массива нераспределенных водотоков с закрытием однопиксельных разрывов фрагментов водотоков, полученных при последовательном затоплении и очисткой структурного рисунка от ошибок дешифрирования («висящих водотоков»). Блок 3 включает поиск точек бифуркации рек и «истоков» водотоков первого порядка с помощью «подсчета соседей» (3.1), морфологическое наращивание (дилатансию) точек бифуркации с помощью дискового «структурного элемента» (3.2), разъединение дешифрированного рисунка в точках бифуркации (3.3), маркирование фрагментов гидросети (3.4) и выбор тех ее фрагментов, которым принадлежат точки начала, в качестве водотоков извлекаемого текущего порядка в матрицу порядков водотоков (МПВ) (3.5). После изъятия водотоков текущего порядка матрица нераспределенных водотоков нуждается в очистке от утолщений в точках бифуркации. Для этого применяется функция 3D скелетизации *skimage.skeletonize_3d* (3.6). После этого текущий порядок водотоков увеличивается на 1 и происходит возврат к шагу 3.1. Цикл суперблока 3 завершается при отсутствии точек «истока», доступных для определения. Нераспределенному массиву водотоков присваивается последний достигнутый в процессе итераций порядок (число Стралера). В завершении реализующая алгоритм программа формирует и записывает в файл изображение порядка водотоков в формате с геопривязкой (GeoTiff).

Результаты исследования и их обсуждение

Применение описанного алгоритма позволило дешифрировать эрозионную сеть с использованием цифровой модели рельефа SRTM для тестового региона. Оценка качества работы модели производилась визуальным сопоставлением полученной крупномасштабной модели гидросети и фактических данных (рис. 2).

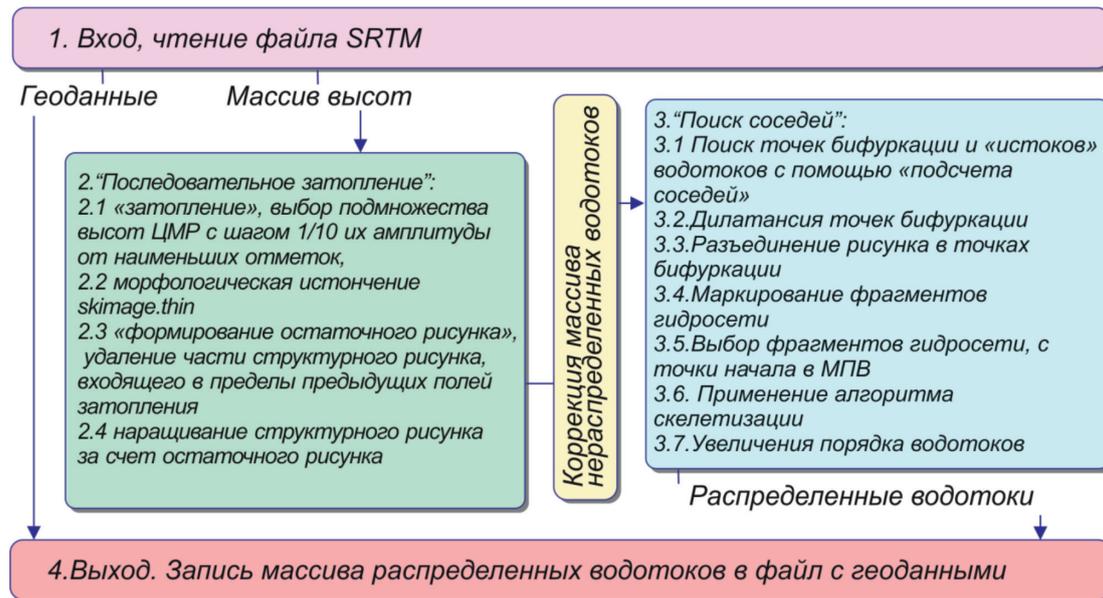


Рис. 1. Блок-схема предлагаемого алгоритма выделения сети водотоков

Результирующая матрица порядков водотоков позволяет охарактеризовать соответствие длин выделенных водотоков и их порядков. Идентифицированные точки бифуркации и «истоки» показаны в виде дополнительного растрового слоя (рис. 2, а). Они позволяют нагляднее представить работу алгоритма. Сопоставление результатов дешифрирования с фактическими данными позволяет говорить о его достаточной достоверности: точность выделения водотоков достигает 83 % процентов (в сопоставлении с глазомерным дешифрированием водотоков на том же участке). Вместе с тем используемый алгоритм не требует применения ресурсоемких коммерческих настольных продуктов, таких как ArcGIS. Интеграция алгоритма и библиотеки GDAL позволит получать в результате работы программы файл с географической привязкой, который можно включить в ГИС-проект.

Для тестового района (пересеченная горно-таежная местность, перепад высот до 997 м) было проведено сопоставление рассматриваемого алгоритма с ближайшим аналогом, дополнением Matlab TopoToolbox [7], результат работы которого показан на рис. 3. Необходимо отметить, что обоими алгоритмами, не требующими пользовательских действий при анализе, были допущены неточности в распознавании водотоков. Так, модуль TopoToolbox, несмо-

тря на высокую точность распознавания водотоков 1-го порядка, допустил разрывы в рисунке гидросети (рис. 3, а, цифры на схеме 1 и 2), что привело к снижению максимального распознанного им порядка (числа Стралера) для эрозионной сети (распознан 4-й, в действительности 5-й порядок). Морфологический алгоритм, предлагаемый нами, установил максимальный распознанный порядок верно, однако им неверно был оконтурен один из притоков (рис. 3, б, цифра на схеме 1), а также объединены в действительности разобщенные элементы гидросети (рис. 3, б, цифра на схеме 2).

Несмотря на то, что оба алгоритма вполне справились с поставленной задачей, для интерпретации результата дешифрирования важно верно определить максимальный порядок водотоков, что не вполне удалось сделать близкому аналогу. Выявленные несовершенства в работе предлагаемого нами алгоритма морфологического дешифрирования послужат основой для дальнейшего развития с целью создания на его основе прикладного программного обеспечения.

Заключение

Рассмотрены особенности и преимущества морфологического анализа изображений для выделения рисунка эрозионной сети и сегментации порядков ее водотоков.

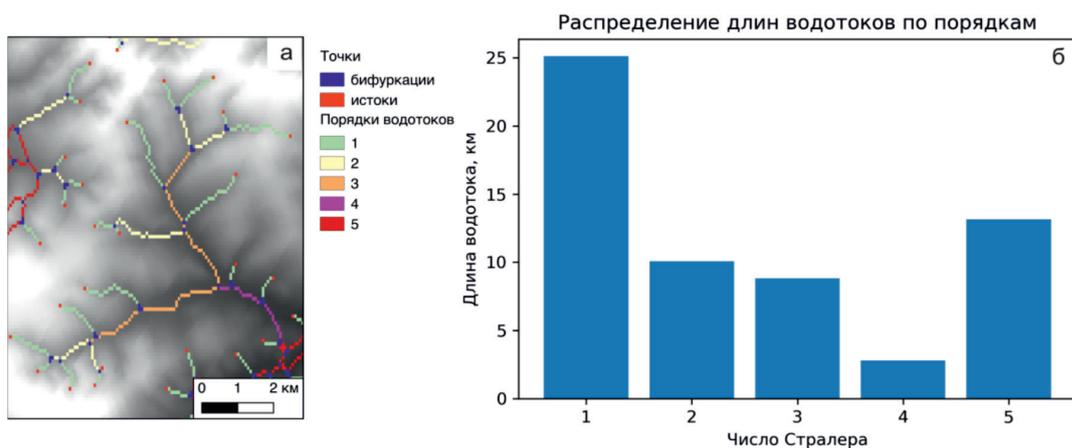


Рис. 2. ЦМР и выявленные с помощью прилагаемого алгоритма водотоки на тестовой площадке (а), диаграмма распределения длин водотоков по их порядкам (числу Стралера) (б)

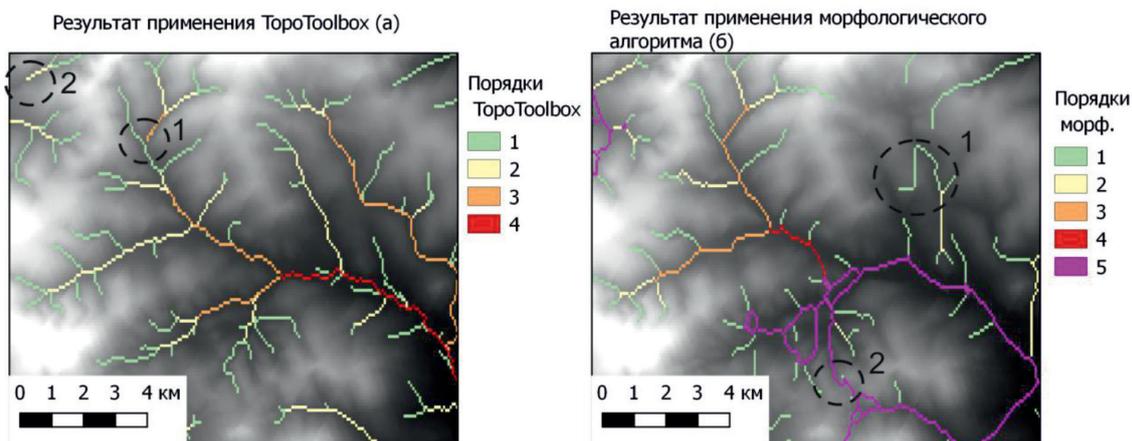


Рис. 3. Сопоставление результатов работы Matlab ToroToolbox (а) и рассматриваемого морфологического алгоритма (б)

Проведенное сопоставление результатов работы предлагаемого алгоритма с ближайшим аналогом, модулем Matlab ToroToolbox [7], выявило недостатки в работе обеих программ, однако разрабатываемый алгоритм корректнее установил максимальный порядок водотоков для территории. С помощью описанного морфологического алгоритма, при его дальнейшем развитии, может быть создано или модернизировано программное обеспечение для географических, геологических и природоохранных целей на основе свободно распространяемых данных ЦМР SRTM. Разработанный алгоритм, опробованный в виде консольного приложения на языке Python, может быть

реализован самостоятельно или в виде дополнения настольной ГИС.

Список литературы / References

1. CGIAR Consortium for Spatial Information. SRTM data [Electronic resource]. URL: <http://srtm.csi.cgiar.org> (date of access: 22.11.2018).
2. NASA. SRTM Data Coverage [Electronic resource]. URL: <https://www2.jpl.nasa.gov/srtm/coverage.html> (date of access: 22.11.2018).
3. Strahler A.N. Quantitative analysis of watershed geomorphology. *Eos Trans, AGU.* 38(6). 1957. P. 913–920.
4. Павлова А.Н. Геоинформационное моделирование речного бассейна по данным спутниковой съемки SRTM (на примере бассейна р. Терешки) // Известия СГУ. Новая серия. Серия Наука о Земле. 2009. Т. 9. № 1. С. 39–44. DOI:10.0000/cyberleninka.ru/article/n/geoinformatsionnoe-modelirovanie-rechnogo-basseyna-po-dannym-sputnikovoy-semki-srtm-na-primere-basseyna-r-tereshki.

Pavlova A.N. Geomodeling of River Basin Based of the Space Data SRTM (as Example of Tereshka River) // *Izvestiya SGU. Novaya seriya. Seriya Nauki o Zemle*. 2009. Т. 9. № 1. P. 39–44 (in Russian).

5. Sumit Das Geomorphic characteristics of a bedrock river inferred from drainage quantification, longitudinal profile, knickzone identification and concavity analysis: a DEM-based study / Das Sumit. *Arabian Journal of Geosciences* (2018) 11:680. DOI: 10.1007/s12517-018-4039-8.

6. Kusák M., Kropáček J., Vilímek V., Schillaci C. Analysis of the influence of tectonics on the evolution of valley networks based on SRTM DEM, Jemma river basin, Ethiopia / M. Kusák, J. Kropáček, V. Vilímek, C. Schillaci. *Geogr. Fis. Dinam. Quat*. 2016. 37–50. DOI: 10.4461/GFDQ.2016.39.4.39.

7. Schwanghart, W., Kuhn, N.J., 2010. TopoToolbox: a set of Matlab functions for topographic analysis. *Environ. Modell. Softw.* 2010. 25. P. 770–781. DOI: 10.1016/j.envsoft.2009.12.002.

8. Bidhyananda Yadav, Kirk Hatfield Stream network conflation with topographic DEMs. *Environmental Modelling & Software*. 2018. V. 102. P. 241–249. DOI:10.1016/j.envsoft.2018.01.009.

9. Xiaoqian Zhao, Yanjun Su, Tianyu Hua, Linhai Chena, Shang Gaoa,b, Rui Wang, Shichao Jin, Qinghua Guo A global corrected SRTM DEM product for vegetated areas. *Remote Sensing Letters*, 2018. V. 9(4). P. 393–402. DOI: 10.1080/2150704X.2018.1425560.

10. Scikit-image. Image processing in Python [Electronic resource]. URL: <https://scikit-image.org> (date of access: 21.11.18).