

# ИСПОЛЬЗОВАНИЕ МЕТОДОВ КОМПЬЮТЕРНОГО ЗРЕНИЯ ДЛЯ АВТОМАТИЗАЦИИ ПРОЦЕССА ОБРАБОТКИ ФОТОГРАФИЙ ГЕОЛОГИЧЕСКИХ ОБРАЗЦОВ

## APPLICATION OF COMPUTER VISION METHODS FOR AUTOMATING THE PROCESSING OF GEOLOGICAL SAMPLE PHOTOGRAPHS

**D. Vorobev  
R. Semendyaev**

*Summary.* This article presents the development of a text recognition system designed for geological samples in geotechnical laboratories. Geotechnical laboratories often encounter a labor-intensive process when generating electronic records for sample codes, as thousands of tests are conducted each month. To address this issue, we have developed an application that integrates image preprocessing techniques with advanced computer vision algorithms. This combined approach enables the creation of compact files, facilitating the detection and recognition of sample codes. Additionally, the system automates the archiving of accurately identified geological sample images on personal computers. As a result, this application significantly reduces the time and effort required by highly skilled personnel for archiving and managing geological sample images, thereby improving overall efficiency.

*Keywords:* automation, image preprocessing, computer vision, geological samples, text recognition.

**Воробьев Данила Вячеславович**

Санкт-Петербургский Политехнический  
университет Петра Великого  
daniila\_vorobev\_02@list.ru

**Семендяев Родион Юрьевич**

кандидат технических наук, научный сотрудник  
отдела инновационных технологий и цифровых систем,  
филиал АО «Россети ТЦ» — СибНИИЭ  
semendyaev@ntcsib.ru

*Аннотация.* Статья посвящена разработке системы распознавания текста на геологических образцах в геотехнических лабораториях. В настоящее время многие лаборатории сталкиваются с трудозатратным процессом создания электронной копии шифра образца (в месяц может проводиться тысячи испытаний). Для решения данной задачи было разработано приложение, в основе которого лежат как методы предварительной обработки изображений, так и алгоритмы машинного зрения. Благодаря комбинированному подходу удалось сделать компактные файлы для последующего детектирования текста шифров, его распознавания и автоматического создания архива корректно идентифицированных изображений геологических образцов на персональном компьютере. Разработанное приложение позволило сократить время высококвалифицированного персонала при решении задачи архивации и хранения изображений геологических образцов. Проведенное тестирование показало высокую точность распознавания текста, что делает разработанное решение эффективным инструментом для автоматизации обработки данных в геотехнических лабораториях.

*Ключевые слова:* автоматизация, предварительная обработка изображений, компьютерное зрение, геологические образцы, распознавание текста.

### Актуальность

В эпоху наступления 4 промышленной революции геотехнические лаборатории имеют проблемы в управлении различными форматами данных и отсутствием единых стандартов для их хранения и обработки [1-3]. Эти трудности приводят к снижению эффективности работы и увеличению вероятности ошибок, например фотографии геологических образцов, полученные в ходе физико-механических испытаний, при переносе на компьютер теряют идентификацию, что затрудняет их сортировку и поиск.

При выборе инструментов для решения задачи итеративного распознавания текста с фотографий и последующего изменения имени файла на распознанный текст, важно учитывать специфику и требования проекта. Это особенно важно, поскольку обработка фотографий является критическим шагом в создании архива,

который является фундаментальной составляющей проекта. Наиболее известные программные продукты для распознавания текста: ABBYY FineReader, Adobe Acrobat Pro и Tesseract не подходят для данной задачи по следующим причинам:

- ABBYY FineReader — коммерческое ПО, что ограничивает его использование в проектах с ограниченным бюджетом. Данное ПО обладает широким функционалом и сложной системой настроек, что избыточно для задачи распознавания шифров образцов и переименования файлов в соответствии с данным шифром.
- Adobe Acrobat Pro ориентирован на работу с PDF, что делает его менее эффективным для обработки фотографий. Высокая стоимость лицензии также ограничивает его доступность для небольших проектов.
- Tesseract требует тщательной предобработки изображений для высокой точности, что может

потребовать дополнительных усилий. Он не предоставляет встроенные инструменты для автоматизации, такие как переименование файлов.

Таким образом, для решения задачи автоматизации распознавания и переименования файлов требуется разработка специализированного приложения, которое обеспечат необходимую гибкость и функциональность.

В данной статье рассматривается использование методов компьютерного зрения для автоматизации обработки фотографий геологических образцов, обсуждаются технические аспекты предварительной обработки фотографий и разработки графического интерфейса пользователя, их практическое применение и влияние на работу лабораторий.

### Необходимость в фотодокументировании геологических образцов

Фотодокументация геологических образцов является важным элементом геотехнических исследований, поскольку она предоставляет визуальные данные, необходимые для всестороннего анализа и интерпретации результатов физико-механических испытаний. Согласно стандартам [4, 5], фотографическая документация должна проводиться до и после испытаний для подтверждения состояния образцов. Эти визуальные данные фиксируют фактическое состояние образцов и служат основой для определения параметров моделирования, которые максимально точно воспроизводят реальные условия.

Для геологов фотографии также важны для подтверждения или уточнения литологического типа образца [6]. Литология позволяет определить механические и физические свойства грунтов, что влияет на выбор методов расчета и проектирования.

Фотографический анализ может выявить особенности, не заметные при визуальном осмотре, такие как микротрещины или включения, которые существенно влияют на поведение материала под нагрузкой.

### Предварительная обработка изображения

Предварительная обработка данных является важным этапом в любом исследовательском проекте, связанном с анализом изображений [7–9]. Этот этап необходим для улучшения качества данных, устранения шума и артефактов, а также для подготовки данных к дальнейшему анализу. Фотография без предварительной обработки представлена на рисунке 1.

Существует множество методов предварительной обработки изображений, которые могут быть использованы в зависимости от конкретных задач и характери-



Рис. 1. Фотография без предварительной обработки

стик данных. Классическими преобразованиями являются: преобразование изображения к оттенкам серого, размытие по Гауссу и пороговое преобразование.

Цветное изображение обычно представлено в формате BGR (Blue, Green, Red), где каждый пиксель имеет три значения, соответствующие интенсивности синего, зеленого и красного цветов. Стандартная фотография геологического образца имеет разрешение 4160x3120 и занимает в памяти 3,2 Мбайт.

Преобразование в оттенки серого заключается в сведении этих трех значений к одному, которое будет представлять яркость пикселя. Для преобразования цветного изображения в оттенки серого согласно стандарту [10] используется следующая формула:

$$Y = 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B$$

Где  $R$ ,  $G$ , и  $B$  — это значения интенсивности красного, зеленого и синего цветов соответственно, а  $Y$  — это значение яркости в оттенках серого.

Результат преобразования изображения к оттенкам серого представлен на рисунке 2.

Гауссовое размытие используется для сглаживания изображения и уменьшения шума [11]. Это преобразова-



Рис. 2. Изображение в оттенках серого

ние полезно перед выполнением таких задач, как обнаружение краев и сегментация. Этот метод основывается на применении ядра Гаусса к каждому пикселю изображения, которое вычисляется в соответствии с функцией Гаусса:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

Где  $G(x, y)$  — значение гауссового фильтра в точке  $(x, y)$ , а  $\sigma$  — стандартное отклонение, определяющее степень размытия.

Формула для вычисления значения пикселя  $I(x, y)$  в результате применения Гауссова размытия определяется следующим образом:

$$I'(x, y) = \frac{1}{2\pi\sigma^2} \sum_{i=-k}^k \sum_{j=-k}^k I(x+i, y+j) \cdot e^{-\frac{x^2+y^2}{2\sigma^2}}$$

Где:  $I(x, y)$  — значение пикселя в исходном изображении,  $I'(x, y)$  — значение пикселя в размытом изображении,  $\sigma$  — стандартное отклонение Гауссова распределения, определяющее степень размытия,  $k$  — радиус ядра Гаусса, который определяет область соседних пикселей, влияющих на результат.

Изображение с применением гауссового размытия представлено на рисунке 3.



Рис. 3. Размытое изображение

Бинаризация основана на сравнении каждого пикселя изображения с заданным пороговым значением [12]. Если значение пикселя выше или равно этому порогу, пиксель присваивается к одному классу (обычно белый), в противном случае — к другому (обычно черный). Это преобразование можно выразить следующим образом:

$$I'(x, y) = \begin{cases} 255, & \text{если } I(x, y) \geq T \\ 0, & \text{если } I(x, y) < T \end{cases}$$

Здесь:  $I(x, y)$  — значение пикселя в исходном изображении в позиции  $(x, y)$ ,  $I'(x, y)$  — значение пикселя в бинаризованном изображении,  $T$  — пороговое значение.

Выбор порогового значения  $T$  является критическим аспектом бинаризации. При съемке с постоянным освещением параметр  $T$  может быть подобран эмпирическим путем, таким образом, чтобы бинаризованное изображение удовлетворяло поставленным требованиям. Бинаризованное изображение представлено на рисунке 4.

Одним из основных методов обнаружения контуров является алгоритм Кэнни [13], который включает в себя несколько ключевых шагов:

1. Гауссово размытие.
2. Операторы Собеля для вычисления градиентов.

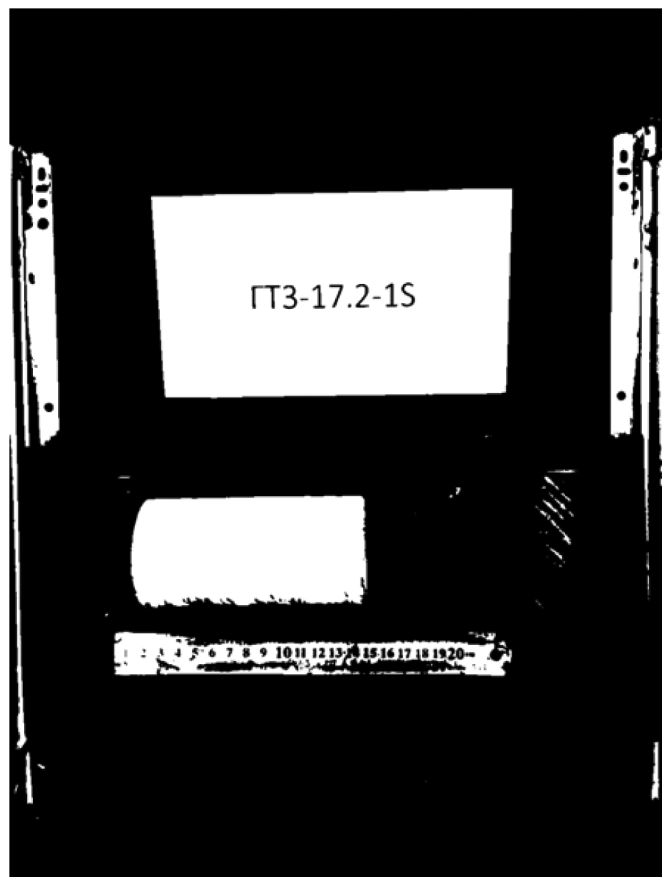


Рис. 4. Бинаризованное изображение

3. Вычисление величины и направления градиента.
4. Подавление не максимумов.
5. Двойная пороговая фильтрация для поиска границ.
6. Трассировка области неоднозначности для отнесения пикселей с промежуточными значениями к границе или их подавлению в противном случае.

Результат сегментации экрана монитора представлен на рисунке 5.

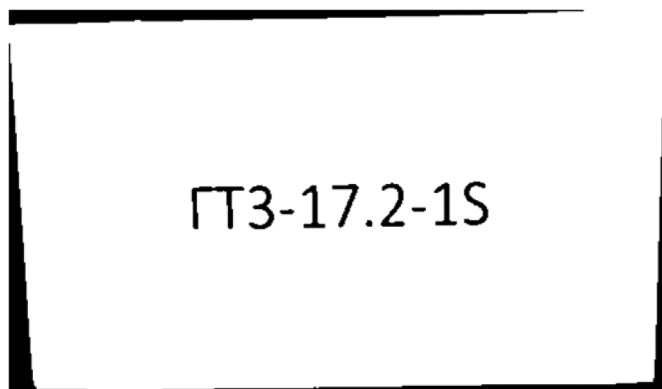


Рис. 5. Вырезанный экран монитора

Получившееся изображение готово к распознаванию. Для извлечения текста с обработанного изо-

бражения, приведенного к бинаризованному виду, используется библиотека «EASY OCR» [14]. В Таблице 1 представлены результаты преобразований изображения, включая изменения веса файла и затраченное время на каждом этапе. Можно заметить, что описанные выше преобразования позволили уменьшить размер файла в 41 раз, что существенно ускоряет извлечение информации с изображения.

Таблица 1.

Эффективность преобразования изображений

Преобразование	Вес изображения, КВ	Затраченное время, с
Изображение без обработки	3221.65	–
Изображение в оттенках серого	2863.38	0.13
Размытое изображение	1691.38	0.12
Черно-белое изображение	271.73	0.07
Сегментированное изображение	77.41	0.01

#### Рабочий процесс разработанного приложения

Разработанное приложение предназначено для упрощения поиска фотографий в папке путем автоматического переименования файлов на основе распознанного текста шифра, отображаемого на мониторе электронного устройства. Разработка архитектуры приложения осуществлялась с общепринятыми подходами и парадигмами, которые освещены в [15–17]. Рабочий процесс разработанного приложения представлена на рисунке 6.

Функционал приложения включает в себя следующие возможности:

1. Использование компьютерного зрения для итеративной обработки фотографий.
2. Распознавание текста шифра на русском и английском языках.
3. Требование постоянного освещения в камере или боксе во время фотографирования для обеспечения высокого качества распознавания текста.
4. Способность распознавать текст шифра, отображаемый на мониторе электронного устройства, независимо от типа шрифта, используемого для его отображения.

Выбор языка программирования Python [18,19] для разработки приложения обусловлен его обширной экосистемой библиотек, таких как PyQt5 [20], OpenCV, EASY OCR, что ускоряет интеграцию функций распознавания текста. Легкость синтаксиса Python облегчает поддержку и развитие кода, а его кроссплатформенность расширяет доступность приложения. Широкая поддержка сообщества и доступ к современным технологиям делают Python оптимальным выбором для задач автоматизации и обработки изображений.

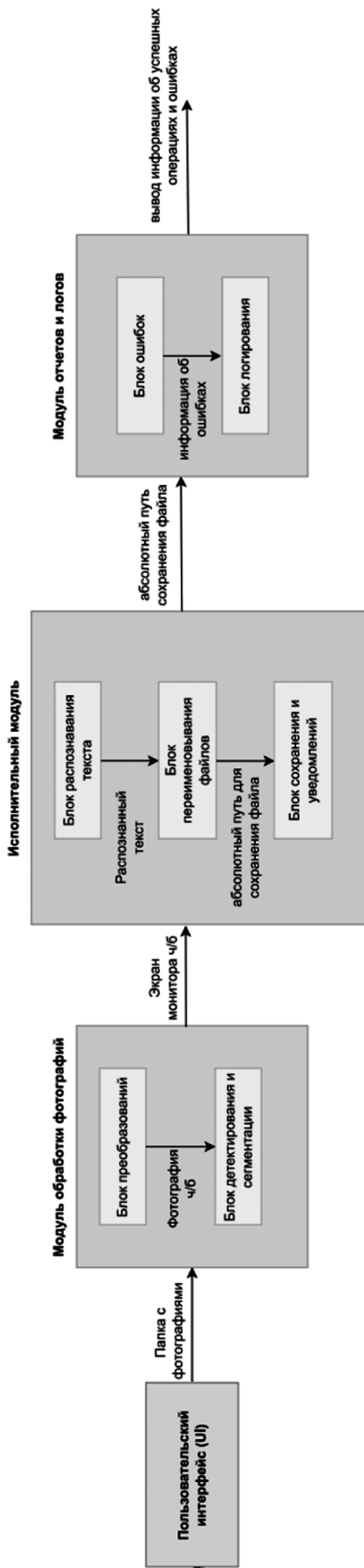


Рис. 6. Рабочий процесс приложения

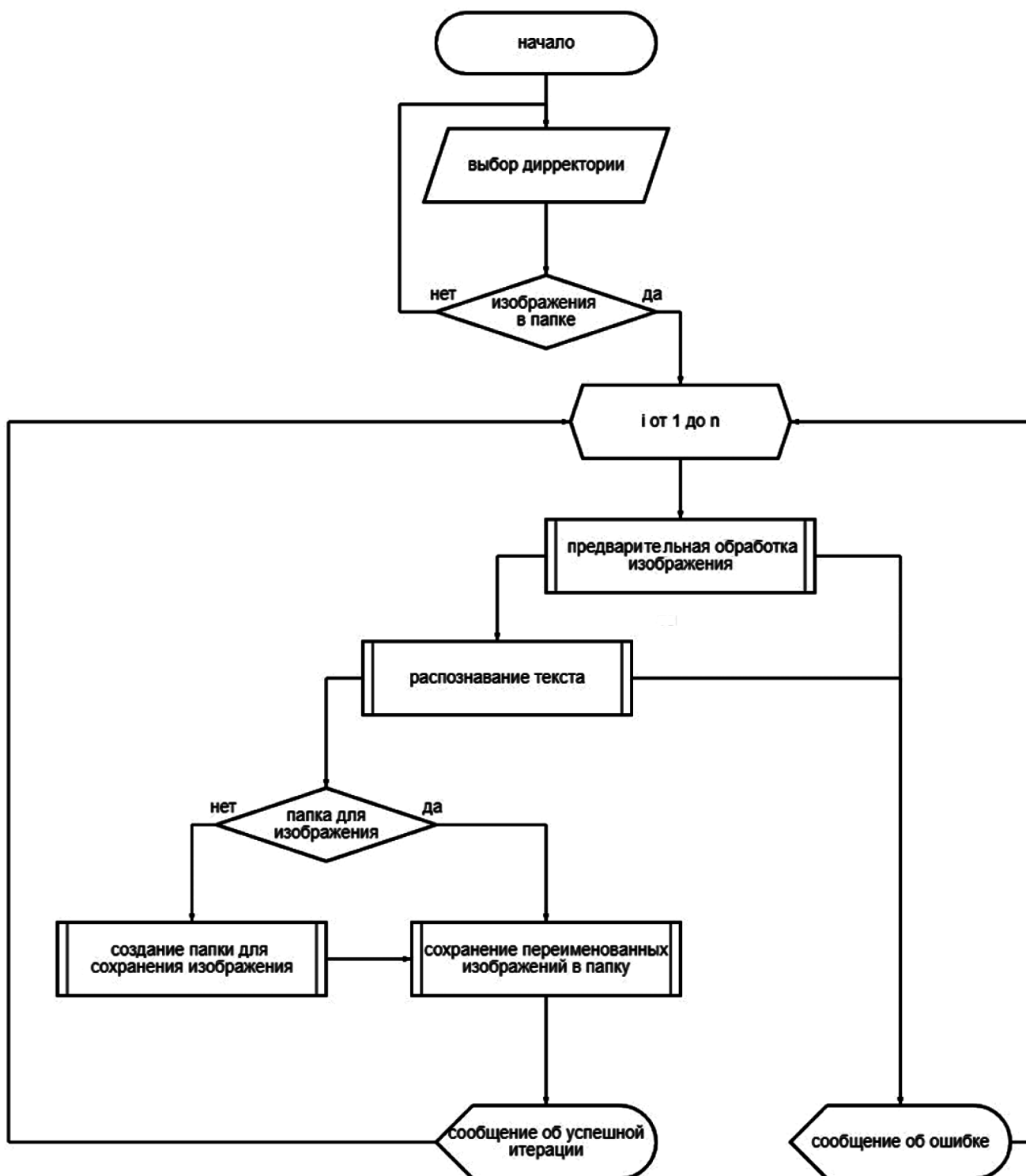


Рис. 7. Блок схема приложения

Существующие фреймворки для разработки программных комплексов имеют разную степень настройки конфигурации приложения. Поскольку разработанная программа не требует от пользователя большого количества операций для запуска обработки фотографий, было принято решение воспользоваться фреймворком PyQt5. Этот фреймворк прост, удобен и позволяет до-

статочно быстро создать приложение, что очень важно в условиях ограниченного времени.

Основными функциональными элементами приложения являются:

1. Кнопка открытия проводника для выбора необходимой директории
2. Кнопка запуска

3. Шкала прогресса для отображения хода выполнения программы
4. Диалоговое окно для вывода сообщений пользователю

Также интерфейс имеет вкладку меню в верхнем левом углу (красная рамка), в котором пользователь может остановить выполнение программы, если есть такая необходимость и вкладка помощь, которая позволяет открыть диалоговое окно со справочной информацией о требованиях и функциональности приложения.

После выбора директории и нажатия кнопки 'ЗАПУСТИТЬ' начинается процесс обработки фотографий. Блок-схема алгоритма представлена на рисунке 7.

### Заключение

В результате работы был проведен анализ методов предварительной обработки изображений, исследо-

ваны возможности различных библиотек компьютера зрения, рассмотрены разнообразные языки программирования и паттерны создания архитектуры приложения. Это позволило разработать приложение на базе библиотеки «EASY OCR», с эффективным (критерий эффективности — быстродействие приложения) процессом переименования файлов и архивации их на персональном компьютере в упорядоченных директориях.

Кроме того, автоматизация обработки фотографий привело к исключению человеческого фактора из этапов обработки изображений, минимизировало затраты времени и ресурсов, что, в свою очередь, привело к повышению эффективности работы и снижению операционных издержек. Таким образом, внедрение данного приложения не только ускоряет обработку фотографий, но и приносит значительный экономический эффект за счет автоматизации рутинных задач.

### ЛИТЕРАТУРА

1. Разводовский Д.Е., Шулятьев С.О., Ставницер Л.П. Применение BIM в геотехнике // *Жилищное строительство*. 2018. № 11. С. 3–8.
2. Hamman ECF, du Plooy DJ, Seery J (2017) Data management and geotechnical models. Proc. Deep Mining 2017: 8th Deep & High Stress Mining Cont. Perth, ACG: 401–487.
3. В. Neil & Reynolds Mark & Nicoll Sam & Bran Dinella. (2018). Geotechnical data management and visualization systems: meeting the data challenge of the 21st century and maximizing value for open pit mines.
4. ASTM D7012-14. Стандартные методы испытаний для сжимающей прочности и модуля упругости породных образцов в режиме одноосного и трёхосного сжатия. — Вест Коншохокен: ASTM International, 2014. — 9 с. — (ASTM International Standard).
5. ГОСТ 21153.2–84. Грунты. Методы лабораторного определения прочности по результатам компрессионных испытаний. — Введ. 01.01.1986. — М.: Изд-во стандартов, 1984. — 9 с.
6. ГОСТ 25100–2011. Грунты. Классификация. — Введ. 01.07.2012. — М.: Стандартиформ, 2012. — 43 с.
7. Фокин Д.С. Предварительная обработка изображений сцены для выделения искомым объектов / Д.С. Фокин, С.Н. Карабцев, Т.С. Рейн // *Международный научно-исследовательский журнал*. — 2020. — №3 (93). — URL: <https://research-journal.org/archive/3-93-2020-march/predvaritelnaya-obrabotka-izobrazhenij-sceny-dlya-vydeleniya-iskomykh-obektov> (дата обращения: 02.10.2024). — doi: 10.23670/IRJ.2020.93.3.003.
8. Хорн Б.К.П. Зрение роботов. — М.: Мир, 1989. — 487 с.
9. Техническое зрение роботов / под ред. А. Пью. — М.: Машиностроение. — 1987. — 320 с.
10. ITU-R Rec. BT.709-6. Параметры вещательного телевидения высокой четкости (HDTV) для студий и международного обмена программами. — Женева: Международный союз электросвязи (ITU), 2015. — 24 с.
11. Vazhitov, Sergey Denisovich. О восстановлении параметра размытия в задаче оптического секционирования. *Современные информационные технологии и ИТ-образование*, [S.l.], v. 18, n. 1, p. 20–27, mar. 2022. ISSN 2411–147
12. Горитов А.Н. Предварительная обработка изображений в системах технического зрения / А.Н. Горитов // *Доклады ТУСУР*. — 2018. — Т. 21, № 4–1. — С. 53–58. DOI: 10.21293/1818–0442–2018-21-4-1-53-58.
13. Гизатуллин З.М. Метод обнаружения контуров на основе весовой модели изображения / З.М. Гизатуллин, С.А. Ляшева, О.Г. Морозов, М.П. Шлеймович // *Компьютерная оптика*. — 2020. — Т. 44, № 3. — С. 393–400. — DOI: 10.18287/2412–6179-CO-615.
14. Jaided M., & Team. (2020). EasyOCR: Ready-to-use OCR with 80+ supported languages. Retrieved from <https://github.com/JaidedAI/EasyOCR>
15. Джонсон Р. Приемы объектно-ориентированного проектирования. Паттерны проектирования. М.: Питер, 2013. 368 с. ISBN: 978-5-459-01720-5
16. Фаулер М. Шаблоны архитектуры корпоративных приложений. Бостон: Addison-Wesley Professional, 2012. 560 с. ISBN: 978-0321127426
17. Цербренко К.Н. Совершенствование архитектуры предприятия на основе функционально-структурного моделирования // *Международный журнал гуманитарных и естественных наук*. 2022. № 7–2. С. 211–214. DOI: 10.24412/2500–1000-2022-7-2-211–214
18. Lutz M. (2013). *Learning Python* (5th ed.). O'Reilly Media.
19. Van Rossum G., Warsaw B., & Coghlan N. (2001). PEP 8 — Style Guide for Python Code. Python Software Foundation. Retrieved from <https://peps.python.org/pep-0008/>
20. Riverbank Computing. (2023). *PyQt5 Documentation*. Retrieved from <https://www.riverbankcomputing.com/software/pyqt/intro>

© Воробьев Данила Вячеславович (danila\_vorobev\_02@list.ru); Семендяев Родион Юрьевич (semendyaev@ntcsib.ru)

Журнал «Современная наука: актуальные проблемы теории и практики»