

ВЕСТНИК НОВОСИБИРСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА

Научный журнал
Основан в ноябре 1999 года

Серия: Информационные технологии

2021. Том 19, № 3

СОДЕРЖАНИЕ

<i>Адов Д. Ю.</i> Разработка программного обеспечения для управления виртуальными объектами при помощи вызванного потенциала P300	5
<i>Белоус И. А., Клюкман М. В., Ковырнев М. В., Пяткова И. А., Сорока Д. Г.</i> Игрофикация прикладных дисциплин технического направления	17
<i>Шабалина Д. Е., Ланчуковская К. С., Лях Т. В., Чайка К. В.</i> Семантическая сегментация изображений в проекте «Duckietown»	26
<i>Загумённых А. А., Наумова В. В., Ерёмченко В. С.</i> Облачный сервис многомерной обработки количественных данных для решения геологических задач	40
<i>Кугаевских А. В.</i> Новая модель нейрона конца линий, применимая в свёрточных нейронных сетях	50
<i>Тихонов Н. И.</i> Методы Paper2vec и Cite2vec для анализа коллекций научных публикаций	61
<i>Флеенко А. С., Демьяненко А. Ф.</i> Разработка методики перехода к технологиям информационного моделирования в инженерных изысканиях (на примере инженерно-экологических изысканий)	70
Информация для авторов	83

VESTNIK

NOVOSIBIRSK STATE UNIVERSITY

Scientific Journal
Since 1999, November
In Russian

Series: Information Technologies

2021. Volume 19, № 3

CONTENTS

<i>Adov D. Yu.</i> Development of Software for Control of Virtual Objects Using Event-Related Potential P300	5
<i>Belous I. A., Klyukman M. V., Kovyrnev M. V., Piatkova I. A., Soroka D. G.</i> Gamification of Applied Technical Disciplines	17
<i>Shabalina D. E., Lanchukovskaya K. S., Liakh T. V., Chaika K. V.</i> Semantic Image Segmentation in “Duckietown”	26
<i>Zagumennov A. A., Naumova V. V., Eremenko V. S.</i> Cloud Service for Multidimensional Processing of Quantitative Data for Solving Geological	40
<i>Kugaevskikh A. V.</i> New End-Stopped Complex Cell Model Applicable in Convolutional Neural Networks	50
<i>Tikhonov N. I.</i> Paper2vec and Cite2vec Methods for Analyzing Collections of Scientific Publications	61
<i>Fleenko A. S., Demianenko A. F.</i> Development of a Methodology for the Transition to Information Modeling Technologies in Engineering Surveys (On the Example of Engineering and Environmental Surveys)	70
Instructions to Contributors	83

Editor in Chief M. M. Lavrentiev

Vice-Editor A. V. Avdeev

Executive Secretary D. P. Iksanova

Editorial Board of the Series

- I. V. Bychkov*, professor, academician (Irkutsk), *B. M. Glinsky*, professor (Novosibirsk)
A. N. Gorban, professor (Lester, GB), *E. P. Gordov*, professor (Tomsk)
B. S. Dobronets, professor (Krasnoyarsk), *A. M. Elizarov*, professor (Kazan)
G. N. Erokhin, professor (Kaliningrad), *A. I. Kamyshnikov*, professor (Khanty-Mansijsk)
G. P. Karev, professor (Maryland, USA), *N. A. Kolchanov*, professor, academician (Novosibirsk)
M. M. Lavrentjev, professor (Novosibirsk), *V. E. Malyshkin*, professor (Novosibirsk)
N. N. Mirenkov, professor (Aizu, Japan), *N. M. Oskorbin*, professor (Barnaul)
D. E. Palchunov, professor (Novosibirsk), *T. Pizansky*, professor (Ljubljana, Slovenia)
V. P. Potapov, professor (Kemerovo), *O. I. Potaturkin*, professor (Novosibirsk)
V. A. Serebryakov, professor (Moscow), *A. V. Starchenko*, professor (Tomsk)
S. I. Smagin, professor, corresponding member of RAS (Khabarovsk)
D. A. Tusupov, professor (Astana, Kazakhstan)
V. V. Shajdurov, professor, corresponding member of RAS (Krasnoyarsk)
Yu. I. Shokin, professor, academician (Novosibirsk)

*The journal is published quarterly in Russian since 1999
by Novosibirsk State University Press*

The address for correspondence

Faculty of Information Technologies, Novosibirsk State University

1 Pirogov Street, Novosibirsk, 630090, Russia

Tel. +7 (383) 363 42 46

E-mail address: inftech@vestnik.nsu.ru

On-line version: <http://elibrary.ru>

Научная статья

УДК 004.415.2.031.43

DOI 10.25205/1818-7900-2021-19-3-5-16

Разработка программного обеспечения для управления виртуальными объектами при помощи вызванного потенциала P300

Дмитрий Юрьевич Адов

Новосибирский национальный исследовательский государственный университет

Новосибирск, Россия

d.adov@g.nsu.ru, <https://orcid.org/0000-0001-5220-2753>

Аннотация

Рассмотрен принцип работы интерфейсов «мозг – компьютер» (ИМК) и метод детектирования фокуса внимания человека, основанный на вызванных потенциалах мозга (P300), проведен обзор существующих аппаратных и программных решений для реализации ИМК, выявлены их преимущества и недостатки. Спроектировано и реализовано программное обеспечение, которое, основываясь на методах работы с вызванными потенциалами, позволяет совершать выбор желаемого стимула из набора предъявляемых.

Для получения электроэнцефалограммы была использована «Система для регистрации восьмиканальной ЭЭГ» от компании «BiTronics Lab», основанная на платформе Arduino. Фильтрация сигнала, обучение классификатора и визуализация реализованы на языке программирования Python.

Ключевые слова

электроэнцефалография, ЭЭГ, программное обеспечение, интерфейс «мозг – компьютер», ИМК, вызванные потенциалы, P300

Благодарности

Автор выражает признательность Ольге Андреевне Джафаровой за идею для работы и профессиональные консультации, а также лаборатории «Инжевика» ФИТ НГУ за предоставленное оборудование

Для цитирования

Адов Д. Ю. Разработка программного обеспечения для управления виртуальными объектами при помощи вызванного потенциала P300 // Вестник НГУ. Серия: Информационные технологии. 2021. Т. 19, № 3. С. 5–16. DOI 10.25205/1818-7900-2021-19-3-5-16

Development of Software for Control of Virtual Objects Using Event-Related Potential P300

Dmitrii Yu. Adov

Novosibirsk State University

Novosibirsk, Russian Federation

d.adov@g.nsu.ru, <https://orcid.org/0000-0001-5220-2753>

Abstract

The article considers the principle of operation of brain-computer interfaces (BCI) and a method for detecting the focus of a person's attention using event-related potential (P300). The review of the existing hardware and software solutions for the implementation of BCIs was performed including the identification of their advantages and disadvantages. The program that allows you to choose the desired stimulus from a variety of presented was developed.

© Адов Д. Ю., 2021

ISSN 1818-7900 (Print). ISSN 2410-0420 (Online)

Вестник НГУ. Серия: Информационные технологии. 2021. Том 19, № 3. С. 5–16

Vestnik NSU. Series: Information Technologies, 2021, vol. 19, no. 3, pp. 5–16

An electroencephalograph of the BiTronics Lab company on the Arduino platform was used to receive the signal. Signal filtering, classifier training and visualization are implemented in Python.

Keywords

electroencephalography (EEG), software, brain-computer interface (BCI), event-related potential, P300

Acknowledgements

The author expresses gratitude to Olga A. Jafarova for the idea for work and professional advice, as well as to the laboratory "Inzhevika" of the FIT NSU for the equipment provided

For citation

Adov D. Yu. Development of Software for Control of Virtual Objects Using Event-Related Potential P300. *Vestnik NSU. Series: Information Technologies*, 2021, vol. 19, no. 3, p. 5–16. (in Russ.) DOI 10.25205/1818-7900-2021-19-3-5-16

Введение

Следующий крупный скачок в развитии человечества, вероятнее всего, будет связан с нейротехнологиями и постепенным увеличением производительности умственного труда за счет интеграции вычислительных машин и мозга человека. Развитие этого направления ускорится после завершения расшифровки работы мозга и в связи с биотехнологической революцией, которая началась после расшифровки генома человека.

Применение интерфейсов «мозг – компьютер» в образовательных целях позволит резко увеличить скорость и объем усвоения новой информации, а развитие нейрогарнитур, модуляции памяти и нейрофитнеса приведет к заметному усилению когнитивных способностей человека¹.

В связи с постоянным ростом рынка текущая ситуация характеризуется дефицитом кадров в области нейротехнологий. Одной из самых разумных форм развития индустрии является сеть «кружков» – современно методически оснащенных центров, которые объединяют специалистов в области информационных технологий, физиологии, кибернетики и робототехники. Основным фокусом работы таких центров должно стать создание учебных программ². Поэтому мной принято решение разработать программное обеспечение для проведения тренингов с использованием биологической обратной связи, в частности вызванный потенциал P300. На основе работы было создано учебно-методическое пособие для развития и продвижения нейротехнологий среди школьников.

Объект работы – программное обеспечение для реализации интерфейса «мозг – компьютер». **Предмет** работы – разработка программного обеспечения, реализующего систему «мозг – компьютер», а также разработка учебно-методического комплекса (УМК), позволяющего познакомиться с нейротехнологиями учащимся школ.

Цель – разработать программное обеспечение для реализации интерфейса «мозг – компьютер», который позволит совершать выбор желаемого стимула из набора предъявляемых, используя классификатор на базе когнитивных компонентов вызванных потенциалов (P300) для дальнейшего создания УМК.

Для достижения этой цели были поставлены следующие **задачи**:

- анализ предметной области;
- ознакомление с существующими решениями, выявление их недостатков и особенностей;
- составление требований для разрабатываемого интерфейса «мозг – компьютер»;
- реализация программной части;
- проведение апробации.

¹ Национальная технологическая инициатива. Нейронет. URL: <https://nti2035.ru/markets/neuronet> (дата обращения 06.03.2021).

² Экспертно-аналитический отчет: Анализ состояния и динамики мирового рынка нейротехнологий. URL: https://www.rvc.ru/upload/iblock/d0d/Analyz_sostoyanoya_i_dinamiki_mirovogo_ryinka_neirotehnologyi.pdf (дата обращения 04.10.2020).

ИМК и вызванные потенциалы P300

В течение многих лет люди предполагали, что электроэнцефалографическая активность мозга может предоставить новый мышечный канал для отправки сообщений и команд во внешний мир – интерфейс «мозг – компьютер» (ИМК). За последние два десятка лет было проведено множество исследовательских программ по отношению к ИМК. В связи с постоянно растущим уровнем понимания деятельности мозга, появлением недорогого высокопроизводительного компьютерного оборудования и растущим осознанием потребностей людей с ограниченными возможностями исследователи ИМК сосредоточены на разработке новых дополнительных технологий коммуникации и контроля для людей с тяжелыми нервно-мышечными расстройствами. К таким можно отнести боковой амиотрофический склероз, инсульт ствола мозга и травму спинного мозга. Одной из основных целей является предоставление частично или полностью парализованным пользователям базовых возможностей общения для коммуникации с окружающими людьми. Более того, подобные системы способны помочь управлять программами обработки текстов или протезами, сопряженными с нейрогабаритурами [1].

ИМК по сути своей – это программно-аппаратные системы детектирования и декодирования различных паттернов электроэнцефалограммы (ЭЭГ), позволяющих осуществлять процедуру биологической обратной связи, например управлять различными механизмами.

Возможности в кодировании намерений человека предоставляют технологии зависимых ИМК. В основе этих ИМК лежит регистрация реакций ЭЭГ на внешнюю стимульную среду, организованную таким образом, чтобы в специфике ответов ЭЭГ можно было определить текущий фокус внимания человека к одному из этих объектов, например к кнопке с соответствующими пиктограммами. Иначе говоря, внешние стимулы начинают кодировать намерения человека, а внимание человека, направленное на некоторый объект, выдает его намерение. Одной из самых удачных систем с практической реализацией зависимых ИМК является спеллер для набора текстов Фарвела и Дончина. Он был разработан еще в конце 1980-х гг., сейчас же исследователи со всего мира совершенствуют различные его модификации: применение цветовых меток целевых стимулов, гибридизация P300-регламентов, использование подвижных стимульных элементов. В ИМК подобного рода используется известное свойство связанных с событиями потенциалов, когда выходящий из ряда (odd-ball) стимул провоцирует появление повышенной позитивности в вызванных потенциалах примерно через 300 мс после предъявления стимула. Эти скачки и служат маркерами фокуса внимания на конкретном объекте из 30–40 находящихся в зрительном поле. Каждый такой объект в стимульной матрице выделяется в случайном порядке подсветкой и таким образом вызывает реакции ЭЭГ. Стоит отметить, что только один объект будет отличаться реакцией ЭЭГ с увеличенным компонентом P300 – это тот, на котором оператор-испытуемый фокусирует свое внимание. Именно так реализуется стимульно-реактивная среда. В ней алгоритмы могут детектировать последовательные «ходы» фокуса внимания оператора от одного объекта к другому, например, они могут детектировать нажатия кнопок виртуальной клавиатуры, пульта управления или другого виртуального устройства ввода информации. Точность детектирования команд в таком ИМК достигает 95–97 % при скорости выбора символов до 12–15 в минуту [2].

Стоит отметить, что впервые о P300 было сообщено более 50 лет назад. Его открытие было вызвано результатом слияния технологических возможностей усреднения сигналов в режиме реального времени и влияния теории информации на психологические исследования. Все последующие результаты показали роль вероятности стимула и релевантности задачи, что послужило основой для его функционального анализа. В большинстве случаев данные были получены при помощи парадигмы «odd-ball». Ее суть заключается в чередовании элементов с низкой вероятностью с элементами с высокой вероятностью, не являющимися целевыми (или «стандартными»). При регистрации с помощью ЭЭГ сигнал P300 проявляется

как положительное отклонение напряжения с задержкой (между стимулом и ответом) примерно от 250 до 500 мс [3]. Именно на этот промежуток будет обращено внимание при анализе сигнала ЭЭГ.

При работе с нейрофизиологической литературой можно столкнуться различными терминами относительно используемого сигнала. Необходимо учитывать, что с момента первого открытия P300 исследования показали, что P300 состоит из двух подкомпонентов – P3 или P3a и классический P300, который с тех пор был переименован в P3b [4].

Обзор существующих программных решений

Experiment – библиотека Python с открытым исходным кодом. Она предназначена для разработки и проведения экспериментов на основе сценариев в режиме реального времени. Работает на Linux, Windows, Mac OS и Android. Разработчики указывают, что библиотекой могут пользоваться непрограммисты, но стоит обратить внимание на то, что для написания сценариев необходимо владеть высоким уровнем навыков программирования ³.

MNE-Python – библиотека Python с открытым исходным кодом. Используется для получения, анализа и визуализации нейрофизиологических данных человека (ЭЭГ, ЭМГ и др.). Написание сценариев требует высокого уровня навыков программирования. Работает на Linux, Windows, Mac OS ⁴.

BioEra – визуальный дизайнер, применяемый для анализа сигналов в реальном времени. Он может взаимодействовать с большим количеством устройств, имеющих возможность потоковой передачи данных на компьютер. Его можно использовать для игр и самостоятельных исследований. Для создания дизайна не нужны навыки программирования, необходимо только понимание сути процесса. Работает на Windows и Android. Бесплатная версия сильно ограничена, например, в ней нельзя сохранить сценарий, если в нем более 15 элементов ⁵.

BioExplorer – это программа для сбора, обработки и отображения биофизических данных в режиме реального времени. Она предназначена для личного использования в сфере развлечений, образования и экспериментов. В подробном описании beta-версии сказано, что программное обеспечение не подходит для использования в медицинских и исследовательских целях, а создано исключительно для личного использования. Работает только на Windows. Мало совместимых устройств ⁶.

BrainBay – приложение для биологической обратной связи, предназначенное для работы с различными усилителями ЭЭГ (включая аппаратные усилители OpenEEG и OpenBCI). Оно поддерживает функции интерфейса «человек – компьютер» и программную платформу NeuroServer для передачи записей в реальном времени через Интернет и локальную сеть. Доступно под Windows ⁷.

EEGLAB – интерактивный набор инструментов Matlab для обработки непрерывных событий, связанных с ЭЭГ, ЭМГ и другими электрофизиологическими данными. Он включает возможность анализа времени, частоты и артефактов. Предоставляет статистику, связанную с событиями, и несколько полезных режимов визуализации испытаний. EEGLAB работает под Linux, Windows и Mac OS. Это одно из лучших решений среди подобных программ ⁸.

EEGMIR – приложение на стадии альфа. В настоящее время почти нет документации. Приложение способно считывать данные из файла или последовательного порта в одном из

³ URL: <https://www.expyriment.org> (дата обращения 04.12.2019).

⁴ URL: <https://martinos.org/mne> (дата обращения 04.12.2019).

⁵ URL: <http://www.bioera.net> (дата обращения 05.09.2020).

⁶ URL: <http://www.cyberevolution.com> (дата обращения 05.09.2020).

⁷ URL: <http://www.shifz.org/brainbay> (дата обращения 05.09.2020).

⁸ URL: <https://seccn.ucsd.edu/eeglab> (дата обращения 05.09.2020).

нескольких форматов и позволяет отображать анализ в виде столбцов. Совместим с Windows и Linux⁹.

E-Prime Extensions for Brain Products – программное обеспечение, предназначенное для оптимизации экспериментального взаимодействия с системами ЭЭГ Brain Products. Позволяет производить запись ЭЭГ и сохранять в файл, устанавливать маркеры на основе стимулов для более простого анализа данных. Для использования требуется E-Prime 3.0, рекордер BrainVision и сервер удаленного управления¹⁰.

Mind-Body Training Tools – программный продукт с биологической обратной связью, предназначенный как для профессионалов (тренеров и терапевтов), так и для домашних пользователей. Он предлагает набор уже готовых приложений для биологической обратной связи и нейробиоуправления, включая тренировки с использованием сердечного ритма, дыхания, ЭМГ и ЭЭГ. Приложения были разработаны с использованием конструктора BioEra, что говорит о наличии зависимостей. Сложный процесс установки, сильно урезанная бесплатная версия. Работает на Windows¹¹.

NeoRec – программа для регистрации ЭЭГ и других нейрофизиологических сигналов для научных или медицинских исследований. Она позволяет записывать данные в файлы различных форматов и потоки данных, но только для последующего анализа и обработки сторонним программным обеспечением. Доступна для Windows и Android¹².

OpenViBE – программная платформа, предназначенная для проектирования, тестирования и использования ИМК. Она подходит для сбора, фильтрации, обработки, классификации и визуализации сигналов мозга в режиме реального времени. Это бесплатное программное обеспечение с открытым исходным кодом, которое работает в операционных системах Windows и Linux. Основными областями применения OpenViBE являются медицина (помощь людям с ограниченными возможностями, БОС в реальном времени, нейробиоуправление), мультимедиа (погружение в виртуальную реальность, использование в видеоиграх) и другие приложения, связанные с ИМК. Пользователи OpenViBE могут быть как программистами, так и людьми, не знакомыми с программированием. Это позволяет использовать платформу врачам, разработчикам видеоигр, исследователям в области обработки сигналов или робототехники [5].

Paradigm Elements for Ports – программа, позволяющая создавать сценарии экспериментов с помощью интуитивно понятного конструктора и интегрировать их с широким спектром систем ЭЭГ и ЭМГ. Есть возможность отправлять триггеры на устройства, используя параллельный, сетевой и последовательный порты. Доступна для Windows¹³.

Rehacor – программный комплекс, содержащий процедуры биологической обратной связи для широкого применения. Например, можно создать сценарий для работы с частотой сердечбиения, дыханием, температурой, электромиограммой, гальванической реакцией¹⁴.

Обзор популярных аппаратных решений

Muse – это устройство ЭЭГ, широко используемое по всему миру. Muse подключается к вашему мобильному устройству через Bluetooth. Обладает четырьмя активными электродами ЭЭГ. Основным применением является медитация и ускорение процесса обучения.

⁹ URL: <http://uazu.net/eegmir> (дата обращения 05.09.2020).

¹⁰ URL: <https://psychology-software-tools.mybigcommerce.com/e-prime-extensions-for-brain-products-3-0> (дата обращения 05.09.2020).

¹¹ URL: <https://www.stressresilientmind.co.uk/mbtt> (дата обращения 07.09.2020).

¹² URL: <https://mks.ru/product/neorec> (дата обращения 07.09.2020).

¹³ URL: <http://www.paradigmexperiments.com> (дата обращения 07.09.2020).

¹⁴ URL: http://medicom-mtd.com/en/study/rehacor/rehacor_study.html (дата обращения 07.09.2020).

Особенности конструкции и количество электродов не позволяют использовать данное устройство для полноценной работы с вызванными потенциалами ¹⁵.

NeuroSky MindWave – Портативная ЭЭГ гарнитура, совместимая с персональными компьютерами и мобильными устройствами. Предоставляет большой выбор приложений для проведения сеансов биологической обратной связи. Первая версия реализована на одном ЭЭГ-электроде, что не позволяет использовать данное устройство для полноценной работы с вызванными потенциалами ¹⁶.

NVX – это усилитель постоянного тока с 24, 32 или 48 монополярными каналами для электродов и 4 вспомогательными биполярными каналами для датчиков NeoSens. Устройство используется в тренировочных системах, клинических и научных исследованиях в составе компьютерной исследовательской системы для кратковременной регистрации электрофизиологических сигналов, в первую очередь ЭЭГ. *NVX* является профессиональным оборудованием с высокой точностью, но и высокой стоимостью. Подобный вариант сложен в использовании при реализации образовательной деятельности с учащимися школ ¹⁷.

«Система для регистрации восьмиканальной ЭЭГ» *BiTronics Lab* – набор, созданный специально для сборки, программирования и регистрации восьмиканальной электроэнцефалограммы мозга в кружках технического творчества и робототехники в рамках дополнительного образования (8–11 классы). Включает в себя восемь одноканальных датчиков ЭЭГ, которые необходимо собрать в единую установку. Для реализации проектов на нем необходимы навыки программирования на Python, а также базовые знания языка C++ и основ машинного обучения ¹⁸.

Проектирование системы и разработка требований

Одним из основных требований к разрабатываемой системе является простота освоения аппаратной и программной части, так как на основе системы разрабатывается учебно-методический материал для преподавания схемотехники, программирования и нейрофизиологии учащимся школ. Почти все перечисленные ранее программы и аппаратные комплексы не позволяют сконструировать доступные и эффективные системы для работы с вызванными потенциалами.

Программные решения можно разделить на три категории.

1. Простые в использовании программы, позволяющие писать сценарии экспериментов, составляя их из готовых блоков. Сценарии можно быстро редактировать, не обладая навыками программирования, однако у таких решений отсутствует возможность создать визуализацию для собственных экспериментов.

2. Библиотеки, используемые в научно-исследовательской деятельности, которые позволяют получать, обрабатывать данные и качественно визуализировать ЭЭГ-сигнал. Для использования подобных решений необходим высокий уровень навыков программирования, что делает невозможным процесс пошагового обучения.

3. Программные платформы, которые содержат в себе лучшие стороны предыдущих двух вариантов. Они просты в написании сценариев, а также легко интегрируются с высокоуровневыми языками программирования, тем самым позволяя создать хорошую визуализацию. Однако, так как они являются лишь связующим звеном между устройством, получающим ЭЭГ-сигнал, и скриптом, реализующим визуализацию, использование подобных платформ лишь оказывает дополнительную нагрузку на вычислительное устройство.

¹⁵ URL: <https://choosemuse.com> (дата обращения 15.11.2019).

¹⁶ URL: <http://neurosky.com/biosensors/eeg-sensor/biosensors> (дата обращения 15.11.2019).

¹⁷ URL: <https://mks.ru/product/nvx> (дата обращения 15.11.2019).

¹⁸ URL: <https://bitronicslab.com/rs8eeg> (дата обращения 15.11.2019).

Аппаратные решения можно разделить на две категории:

1) устройства с фиксированным положением электродов, не позволяющие вести многоканальную запись сигнала с мозга, что делает невозможным работу с вызванными потенциалами;

2) профессиональные электроэнцефалографы, стоимость которых делает невозможным закупку подобного оборудования в школы.

Для разработки УМК мной принято решение использовать восьмиканальную систему для регистрации ЭЭГ от компании «BiTronics». Набор специально разрабатывался для преподавания нейротехнологий учащимся школ, в том числе для работы с вызванными потенциалами. Поэтому с точки зрения аппаратной части он полностью соответствует требованиям:

- позволяет проводить многоканальную запись электроэнцефалограммы;
- имеет в комплекте шапочки, позволяющие изменять расположение электродов;
- работает на Arduino, что позволяет перехватывать сигнал через поток и обрабатывать и визуализировать его любыми удобными методами;
- совместима с операционными системами Windows и Linux;
- полностью безопасна для использования учащимися школ;
- проста в сборке;
- имеет подробную документацию.

Разработка же программного обеспечения для реализации ИМК, позволяющего совершать выбор желаемого стимула из набора предъявляемых, используя классификатор на базе когнитивных компонентов вызванных потенциалов (P300), является основной целью работы. Для этого необходимо получить сигнал с электроэнцефалографа, предобработать его, обучить классификатор и написать визуализацию для предъявления стимулов. Для программного обеспечения, учитывая то, что на его основе будет разработан образовательный курс, где учащимся 9–11 классов будет предложено самостоятельно разработать подобную систему, основным требованием является простота реализации. Под этим подразумевается использование библиотек с доступной документацией, а также игровая форма представления стимулов.

Сбор установки

Считывание сигнала реализовано с помощью датчиков ЭЭГ «BiTronics Lab-EEG», входящих в комплект «Система для регистрации восьмиканальной ЭЭГ» «BiTronics Lab». Датчики разработаны с использованием Arduino. Arduino – это электронный конструктор и платформа с открытым исходным кодом, основанная на простом в использовании аппаратном и программном обеспечении, созданная для быстрой разработки электронных устройств для новичков и профессионалов. Из-за простого и доступного пользовательского интерфейса Arduino используется при создании различных проектов и приложений. Хотя программное обеспечение Arduino простое в использовании, оно обладает рядом преимуществ для опытных пользователей. Платформа работает на Windows, Linux и Mac. Учащиеся школ, студенты, преподаватели и научные сотрудники используют ее для создания дешевых инструментов для доказательства принципов физики или для начала работы с программированием и схемотехникой. Устройство программируется через идущий в комплекте кабель USB и не требует использования отдельных программаторов¹⁹.

Регистрация сигнала осуществляется с помощью электродов, которые закреплены в шапочке, надетой на человека. Для улучшения электрического контакта между проводящей частью электрода и поверхностью кожи головы требуется нанести электропроводящий гель.

Поскольку модуль ЭЭГ является одноканальным и дифференциальным, потребуется расположить 2 электрода в той области головы, где требуется осуществить считывание сигнала,

¹⁹ URL: <http://arduino.ru/> (дата обращения 08.03.2021).

а рядом с ними расположить референсный электрод. Таким образом, для регистрации всех восьми каналов ЭЭГ я закрепил в шапочке и подключил к модулям 16 сигнальных электродов и 8 референсных, всего 24 электрода. На выходе модуля ЭЭГ мы получаем усиленный аналоговый сигнал ЭЭГ, который необходимо будет оцифровать для дальнейшей обработки.

Оцифровка была осуществлена средствами контроллера Arduino Mega, который позволяет одновременно оцифровывать 8 аналоговых сигналов (максимальное количество – 16). Таким образом, сборка установки сводится к подключению электродной системы (совокупность электродов ЭЭГ с проводами) к модулям ЭЭГ (рис. 1), которые, в свою очередь, подключаются к Arduino Mega. Оцифрованный сигнал с Arduino Mega через USB-кабель подается на персональный компьютер.

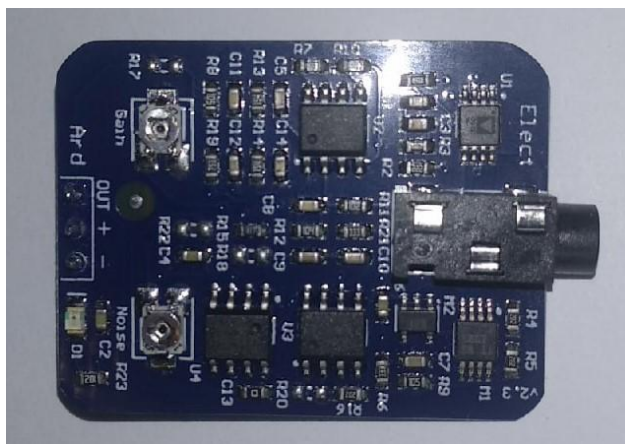


Рис. 1. Одноканальный модуль ЭЭГ
Fig. 1. EEG module for one channel

Ранее соединение модулей с Arduino Mega происходило при помощи макетной платы и большого количества дополнительных проводов типа «папа-папа» (рис. 2).

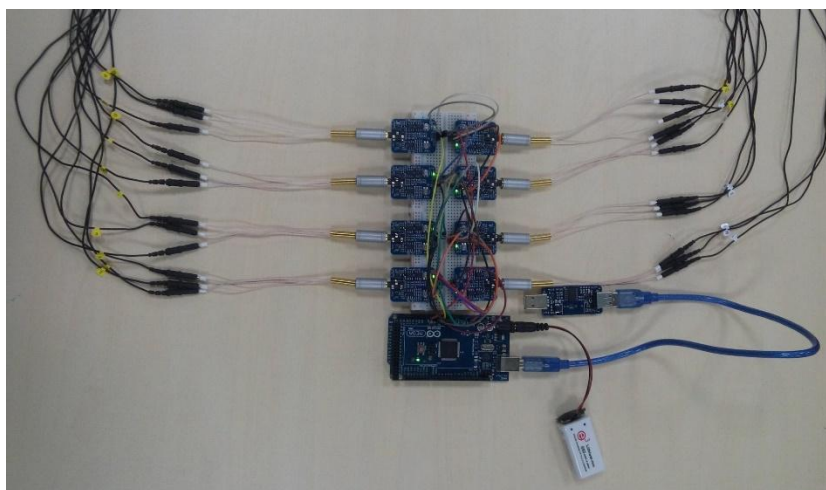


Рис. 2. Общий вид собранной установки
Fig. 2. General view of the assembled installation

В 2020 г. в ЛабИФИТ, на базе которой выполнялась работа, поступили новые наборы, которые содержат в себе плату-переходник (рис. 3). Отсутствие проводов очень упростило процесс сборки, а также позволило избавиться от лишних помех, которые могут возникнуть при малейшем движении. Более подробную инструкцию по сборке оборудования можно найти в комплекте с ресурсным набором «BiTronics Lab-EEG». Стоит отметить, что мной из конструкции был удален фоторезистор, используемый для синхронизации программной и аппаратной частей, так как современные вычислительные машины обладают достаточной мощностью для обработки сигнала в режиме реального времени без дополнительной синхронизации.

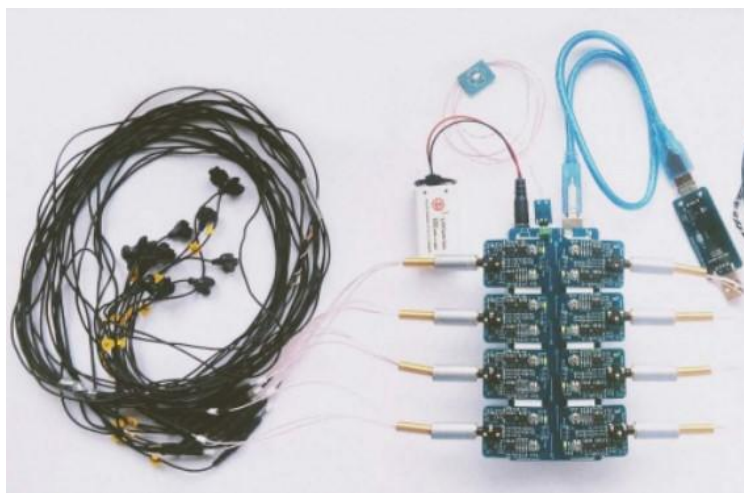


Рис. 3. Общий вид собранной установки с платой-переходником
Fig. 3. General view of the assembled installation to the adapter board

Написание программной части

В первую очередь необходимо установить и настроить среду Arduino IDE. Скачать программное обеспечение можно с официального сайта. Для работы с платами необходимо дополнительно установить библиотеку TimerOne²⁰. Ее использование позволяет вызывать функцию каждые n мс., где n является аргументом.

Для получения сигнала с модулей на плату Arduino был написан скетч, который считывает данные с аналоговых портов и передает их в последовательный порт. Далее необходимо было построить графики сырого сигнала ЭЭГ для детектирования помех и переустановки электродов, если сигнал слишком слабый. После того, как сырой сигнал получен, была задача научиться выделять предполагаемые вызванные потенциалы путем усреднения нескольких эпох и разметки данных, выполняется это при помощи языка программирования Python²¹.

По сути, необходимо работать всего с двумя массивами данных, в одном из которых есть P300 (рис. 4), а в другом нет. Чтобы повысить качество работы классификатора, данные предварительно очищались при помощи фильтра Баттерворта²².

Сама по себе задача классификации направлена на определение типа объекта из двух и более классов. Классификатор – система, которая и определяет тип объекта. Одна из самых

²⁰ URL: <http://bitronicslab.com/neuromodelist#!/tab/242108447-2> (дата обращения 04.04.2019).

²¹ URL: <https://www.python.org> (дата обращения 24.12.2020).

²² URL: https://en.wikipedia.org/wiki/Butterworth_filter (дата обращения 11.02.2021).

популярных библиотек для машинного обучения на Python – Scikit-Learn, в ней содержится также несколько основных типов классификаторов. Для предсказания наличия вызванного потенциала в определенной эпохе я решил использовать классификатор, основанный на методе опорных векторов²³, поскольку этот метод является одним из самых простых и эффективных для решения подобных задач.

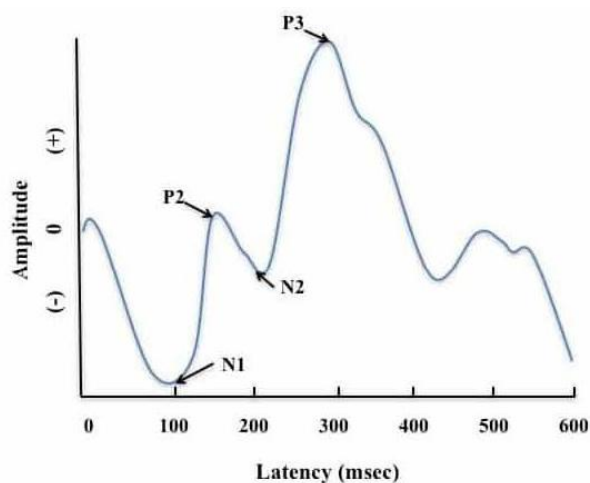


Рис. 4. Вызванный потенциал P300

Fig. 4. Event-related potential P300

Для сбора данных / представления стимула мной было решено использовать лабиринт, в котором игроку предлагается выбрать направление движения. Четыре варианта движения (вверх, вниз, влево, вправо) поочередно подсвечиваются, а когда испытуемый (он же игрок) видит, что загаданное им направление загорается, программа фиксирует вызванный потенциал, после чего игровой персонаж перемещается на экране (рис. 5). Визуализация написана средствами библиотеки PyGame.

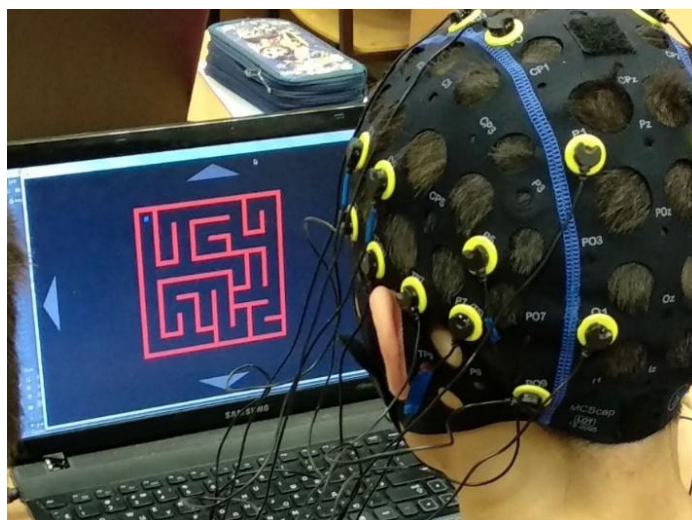


Рис. 5. Выбор пути персонажа при помощи классификации вызванного потенциала

Fig. 5. Choosing a character path using the event-related potential classification

²³ URL: <https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html> (дата обращения 02.02.2021).

Заключение

В работе был проведен анализ предметной области, разработаны требования к системе, спроектировано и реализовано программное обеспечение. Основываясь на методах работы с вызванными потенциалами, разработанная программа позволяет совершать выбор желаемого стимула из набора предъявляемых, используя классификатор на базе когнитивных компонентов вызванных потенциалов (P300). Все поставленные задачи были выполнены, цель достигнута.

Для получения сигнала использована «Система для регистрации восьмиканальной ЭЭГ» от компании «ViTronics Lab», основанная на платформе Arduino. Фильтрация сигнала, обучение классификатора и визуализация реализованы на языке программирования Python.

Разработанная программа была использована при построении образовательного курса по нейротехнологиям в лаборатории «Инжевика» ФИТ НГУ.

Список литературы

1. **Wolpaw J. R., Birbaumer N., McFarland D. J., Pfurtscheller G., Vaughan T. M.** Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, 2002, vol. 113, no. 6, pp. 767–791. DOI 10.1016/S1388-2457(02)00057-3
2. **Каплан А. Я., Кочетова А. Г., Шишкин С. Л., Басюл И. А., Ганин И. П., Васильев А. Н., Либуркина С. П.** Экспериментально-теоретические основания и практические реализации технологии «интерфейс мозг – компьютер» // Бюллетень сибирской медицины. 2013. Т. 12, № 2. С. 21–29. DOI 10.20538/1682-0363-2013-2-21-29
3. **Polich J.** Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 2007, vol. 118, no. 10, pp. 2128–2148. DOI 10.1016/j.clinph.2007.04.019
4. **Squires N. K., Squires K. C., Hillyard S. A.** Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and Clinical Neurophysiology*, 1975, vol. 38, no. 4, pp. 387–401. DOI 10.1016/0013-4694(75)90263-1
5. **Renard Y., Lotte F., Gibert G., Congedo M., Maby E., Delannoy V., Bertrand O., Lécuyer A.** OpenViBE: An Open-Source Software Platform to Design, Test and Use Brain-Computer Interfaces in Real and Virtual Environments. *Massachusetts Institute of Technology Press (MIT Press)*, 2010, vol. 19, no. 1, pp. 35–53. DOI 10.1162/pres.19.1.35

References

1. **Wolpaw J. R., Birbaumer N., McFarland D. J., Pfurtscheller G., Vaughan T. M.** Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, 2002, vol. 113, no. 6, pp. 767–791. DOI 10.1016/S1388-2457(02)00057-3
2. **Kaplan A. Ya., Kochetova A. G., Shishkin S. L., Basyul I. A., Ganin I. P., Vasilev A. N., Liburkina S. P.** Experimental and theoretical foundations and practical implementation of technology brain-computer interface. *Bulletin of Siberian Medicine*, 2013, vol. 12, no. 2, pp. 21–29. (in Russ.) DOI 10.20538/1682-0363-2013-2-21-29
3. **Polich J.** Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 2007, vol. 118, no. 10, pp. 2128–2148. DOI 10.1016/j.clinph.2007.04.019
4. **Squires N. K., Squires K. C., Hillyard S. A.** Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and Clinical Neurophysiology*, 1975, vol. 38, no. 4, pp. 387–401. DOI 10.1016/0013-4694(75)90263-1
5. **Renard Y., Lotte F., Gibert G., Congedo M., Maby E., Delannoy V., Bertrand O., Lécuyer A.** OpenViBE: An Open-Source Software Platform to Design, Test and Use Brain-Computer Interfaces in Real and Virtual Environments. *Massachusetts Institute of Technology Press (MIT Press)*, 2010, vol. 19, no. 1, pp. 35–53. DOI 10.1162/pres.19.1.35

Информация об авторе

Дмитрий Юрьевич Адов, магистрант второго года обучения

Information about the Author

Dmitrii Yu. Adov, 2nd year master student

*Статья поступила в редакцию 20.06.2021;
одобрена после рецензирования 01.08.2021; принята к публикации 01.08.2021
The article was submitted 20.06.2021;
approved after reviewing 01.08.2021; accepted for publication 01.08.2021*

Научная статья

УДК 37.022

DOI 10.25205/1818-7900-2021-19-3-17-25

Игрофикация прикладных дисциплин технического направления

Игорь Александрович Белоус¹
Михаил Валерьевич Ковырнёв²
Михаил Владимирович Клюкман³
Ирина Александровна Пяткова⁴
Дмитрий Геннадьевич Сорока⁵

¹⁻⁵ Владивостокский государственный университет экономики и сервиса
Владивосток, Россия

¹ Igor.Belous@vvsu.ru, <https://orcid.org/0000-0001-8519-4390>

² misha.kovyrnev@gmail.com, <https://orcid.org/0000-0003-3658-9526>

³ misha.supker@gmail.com, <https://orcid.org/0000-0002-0227-8765>

⁴ mellkiss@outlook.com, <https://orcid.org/0000-0002-5657-6282>

⁵ dima.soroka.2014@mail.ru, <https://orcid.org/0000-0001-6656-0682>

Аннотация

Рассмотрены получившие широкое распространение три формы взаимодействия преподавателя и студентов. Основной целью работы является изучение принципиальных отличий между играми и игрофикациями для выделения и модификации наиболее эффективных методик внедрения активных элементов обучения в образовательный процесс. При проведении декомпозиции игр и игрофикации были выявлены особенности активностей. Отмечены различия между целями игр, где нужно победить, и игрофикацией, где нужно научиться. Показано, что выделение оптимальных игровых механик для улучшения процесса обучения, направленных на студентов технических специальностей, происходит по субъективным критериям, а результат – по объективным (успеваемость). Установлено, что применение игровых сценариев при проведении лабораторного практикума объективно повышает успеваемость, тем самым увеличивая сформированность умений и навыков профессиональных компетенций.

Ключевые слова

электронное обучение, методы обучения, игрофикация, сценарий игрофикации, интерактивный метод, активный метод

Для цитирования

Белоус И. А., Клюкман М. В., Ковырнев М. В., Пяткова И. А., Сорока Д. Г. Игрофикация прикладных дисциплин технического направления // Вестник НГУ. Серия: Информационные технологии. 2021. Т. 19, № 3. С. 17–25. DOI 10.25205/1818-7900-2021-19-3-17-25

© Белоус И. А., Клюкман М. В., Ковырнев М. В.,
Пяткова И. А., Сорока Д. Г., 2021

Gamification of Applied Technical Disciplines

Igor A. Belous¹, Mikhail V. Kovyrnev²
Mikhail V. Klyukman³, Irina A. Piatkova⁴
Dmitriy G. Soroka⁵

¹⁻⁵ Vladivostok State University of Economic and Service
Vladivostok, Russian Federation

¹ Igor.Belous@vvsu.ru, <https://orcid.org/0000-0001-8519-4390>

² misha.kovyrnev@gmail.com, <https://orcid.org/0000-0003-3658-9526>

³ misha.supker@gmail.com, <https://orcid.org/0000-0002-0227-8765>

⁴ mellkiss@outlook.com, <https://orcid.org/0000-0002-5657-6282>

⁵ dima.soroka.2014@mail.ru, <https://orcid.org/0000-0001-6656-0682>

Abstract

Three forms of interaction between teacher and students, which have become widespread, are described. The main purpose of the work is to study the fundamental differences between games and gamification in order to identify and modify the most effective methods for introducing active learning elements in the educational process. After decomposition of games and gamification, the following features of activities were revealed. Already at this stage, differences are noticeable between the goals of games where you need to win, and gamification – to learn. It is shown that the selection of optimal game mechanics, to improve the learning process, was carried out according to subjective criteria, and the result according to objective – academic performance, aimed at students of technical specialties. Thus, repeated studies for groups of other specialties can give different, but no less interesting situations. It was found that the use of game scenarios during a laboratory practice objectively increases academic performance, thereby increasing the formation of skills and abilities of professional competencies.

Keywords

e-learning, teaching methods, gamification, gamification scenario, interactive method, active method

For citation

Belous I. A., Klyukman M. V., Kovyrnev M. V., Piatkova I. A., Soroka D. G. Gamification of Applied Technical Disciplines. *Vestnik NSU. Series: Information Technologies*, 2021, vol. 19, no. 3, p. 17–25. (in Russ.) DOI 10.25205/1818-7900-2021-19-3-17-25

Введение

В современном образовании широкое распространение получили три метода взаимодействия преподавателя и студентов [1]: пассивное; активное; интерактивное.

Пассивный метод – это форма взаимодействия преподавателя и студента, в которой преподаватель является основным действующим лицом, управляющим ходом занятия, а студенты выступают в роли пассивных слушателей, подчиненных директивам преподавателя [2].

Активный метод – это форма взаимодействия студентов и преподавателя, при которой они взаимодействуют друг с другом в ходе занятия, и студенты здесь не пассивные слушатели, а активные участники, студенты и преподаватель находятся на равных правах [2].

Интерактивный метод. Интерактивный (Inter – взаимный, act – действовать) означает взаимодействовать, находиться в режиме беседы, диалога с кем-либо. Другими словами, в отличие от активных методов, интерактивные ориентированы на более широкое взаимодействие студентов не только с преподавателем, но и друг с другом и на доминирование активности студентов в процессе обучения [2].

К основным интерактивным формам обучения можно отнести следующие [3]:

- творческие задания;
- работа в группах;
- обучающие игры;
- интерактивная лекция;
- социальные проекты и другие внеаудиторные методы обучения (соревнования, интервью, фильмы, спектакли, выставки);

- тестирование;
- дистанционное обучение;
- разрешение проблем;
- тренинги и др.

В настоящее время перечисленные выше методы обучения успешно реализуются с использованием электронного обучения [4]. Электронное обучение представляет собой одну из форм образовательного процесса, отличающуюся непосредственным и систематическим применением современных информационных технологий и вычислительной техники. Данное понятие в повседневной жизни также может быть заменено такими терминами, как дистанционное, виртуальное, компьютерное, мультимедийное, веб-ориентированное образование и др. Современное состояние средств вычислительной техники и обширное распространение сети Интернет предоставляют возможность для реализации многочисленных преимуществ электронной технологии обучения: удаленность, массовость, высокий уровень интерактивности, обеспечение доступа к электронным библиотекам, формирование единой образовательной среды и т. п. [5].

Основной целью работы является изучение принципиальных отличий между играми и игрофикациями для выделения наиболее эффективных методик внедрения активных элементов обучения в образовательный процесс.

Проведя декомпозицию игр и игрофикаций (геймификаций) были выявлены следующие особенности активностей. Уже на этом этапе заметны различия между целями игр, где нужно победить, и игрофикации, где нужно научиться.

Задачи игрофикации и игр

Игрофикация:

- удержание и концентрация внимания на важных аспектах процессов для снижения эффекта рутинности операций;
- повышение производительности труда.

Игры:

- через игру разобраться с организацией работы и изменить процессы;
- создать условия для достижения цели.

Когда дело доходит до постановки задачи, то тут также видны кардинальные отличия. Если игрофикация направлена на снижение эффекта рутинности без отрыва от основного процесса, то игры самостоятельно ставят участнику задачу, к которой он стремится, хотя сам процесс достижения также может быть монотонным. Главная особенность игрофикации в том, что своей механикой она принципиально отличается от процесса игры направленно-стью на учебный процесс. При более детальном анализе теоретического материала можно вычлениить основные направления игр и игрофикаций и провести сравнительный анализ (табл. 1).

Ключевым моментом в разработке игровых механик стало максимальное воздержание от соревновательных элементов, поскольку студенту необходимо концентрироваться на получении знаний и практических навыков, а не на попытке выполнить задание быстрее конкурентной команды. Концентрация на учебном процессе позволяет конвертировать результаты обучения в практические навыки, которые не только напрямую связаны с его профессиональной деятельностью, но и так называемые «Soft skills», включающих в себя умение быстро адаптироваться к новым условиям работы, налаживать коммуникацию внутри команды, а также быстро оперировать имеющимися знаниями.

Для решения главной задачи работы необходимо создание сценариев, которые позволят полностью отразить суть игрофикации. После анализа типичных активностей при реализации игрофикации в обучении с учетом технического направления подготовки был разработан прототип для внедрения активности в учебный процесс в рамках обозначенных ранее – «Ка-

русель», которая предусматривает несколько сценариев реализации. Его главное направление – работа группы на стендах, однако его можно адаптировать под иные задачи. Данная активность подразумевает распределение членов одной команды по изолированным ролям, т. е. каждый учащийся отвечает за определенный вид работы на время, ограниченное преподавателем, после чего, по сигналу, учащиеся передают свою роль следующему в команде, а на себя берут задачи предыдущего (рис. 1).

Таблица 1

Сравнительный анализ игр и геймификаций по базовым механикам

Table 1

Compared analysis of games and gamification based on common mechanics

Признаки и критерии	Игра	Геймификации
Добровольность участия	Да	Да
Состояние процесса	Автономность, в отрыве от процесса	Интегрированность в деятельность безотрывно от процесса
Правила	Развернутые и иногда сложные	Краткие и простые
Границы мира	Фиксированные	Открытые
Механики	Есть	Есть
Влияние на неигровой результат	Может влиять, а может и не влиять	Не влияет
Конвертируемость результата	Нет	Да
Вознаграждение	Игровая победа	Реальные предметы
Цена проигрыша / выигрыша	Критична	Не критична
Фокусировка	На задаче	На процессе



Рис. 1. Выполнение лабораторных работ по ролям

Fig. 1. Performing a laboratory works by roles

Разделение ролей очень важно для более эффективной работы, поэтому мы прибегнем и к нему (рис. 2). Наш концепт подразумевает одни и те же роли для каждого сценария, а именно:

1) студент, сидящий за стендом и выполняющий основные действия по лабораторной работе;

- 2) студент, заполняющий данные в электронных таблицах, а также делающий скриншоты необходимые для отчета;
- 3) студент, проводящий оформление отчета.



1-я роль.
На стенде (сбор
рабочих схем)



2-я роль.
Фиксация
измерений с
использованием
специального
ПО.



3-я роль.
Заполнение
отчетных
форм.

Рис. 2. Роли в команде «Карусель»
Fig. 2. Roles in team "Carousel"

На данном этапе мы рассматриваем роли для трех человек в команде «Карусели», но количество может свободно меняться для больших групп и достигать пяти человек. Дальнейшее увеличение участников невозможно, так как это будет вносить неудобства в реализацию концепции. Также можно и уменьшать количество до двух человек, но это также вносит серьезные неудобства, поэтому было выбрано оптимальное количество – три. Для наглядного рассмотрения расстановки студентов в группах, мы взяли рисунок лабораторного кабинета ВГУЭС (рис. 3), где в основном проводятся практические занятия для студентов технических специальностей.

Разработаны две концепции сценария развития для «Карусели»: «Карусель внутрикомандная» и «Карусель контрольная».

При реализации первой концепции студенты делятся на группы и распределяют роли, как показано на рис. 2. Далее выполняется лабораторная работа, где прописаны пункты «Смена ролей». В определенный заранее момент студенты одной команды по часовой стрелке меняются местами. Это помогает им в полном объеме поучаствовать в выполнении каждой части лабораторной работы, а также трезво мыслить в условиях стресса и смены места, что зачастую очень сложно.

Так же как и при реализации первой концепции, в концепции «Карусель контрольная» в ходе лабораторной работы обучающиеся встретят опознавательные знаки того, что им нужно поменяться, но здесь будет написан номер студентов, а на доске – время, когда они должны это сделать. Другими словами, запускается тайминг, например, на 15 минут, а в лабораторной работе есть пункт, подписанный как n -я контрольная точка. По истечении времени студенты, чей номер совпал с указанным, должны по схеме поменяться местами на какое-то контрольное время (устанавливается преподавателем), проверить другую команду, помочь им, если у них что-то не получается, или же, наоборот, чему-то научиться и вернуться на

свое место. В отличие от первой концепции, данная требует более тщательной подготовки (рис. 4).

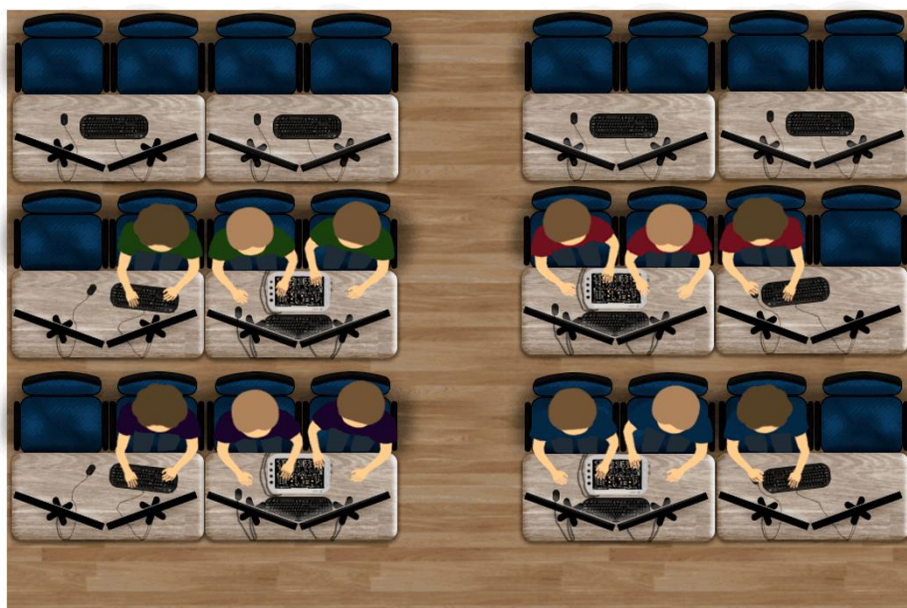


Рис. 3. Кабинет для «Карусели»

Fig. 3. Cabinet for "Carousel"

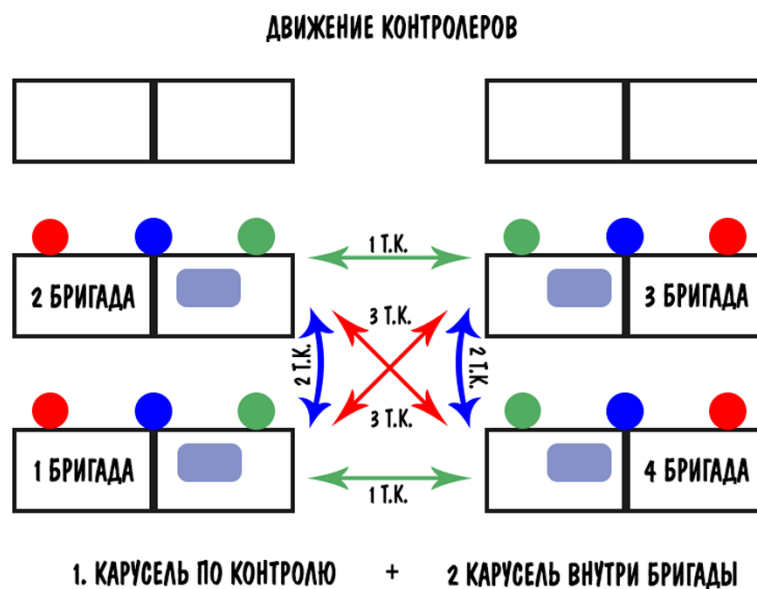


Рис. 4. Схема для «Карусели внутрикомандной»

Fig. 4. Scheme for "Intra-team carousel"

В качестве экспериментальной были выбраны группа БИК-18 и дисциплина третьего семестра обучения бакалавров «Теория сигналов». По итогу курса был вычислен средний балл студента в группе, а также проведено прямое сравнение с результатами групп БИК-16 и БИК-17, которые изучали предмет «классическим методом» (табл. 2, рис. 5).

Таблица 2

Данные по успеваемости до и после внедрения геймификации
в учебный процесс

Table 2

Performance data before and after the introduction of gamification
in the educational process

Группа	Количество человек в группе	Степень усвоения материала			
		не усвоен	усвоен удовлетворительно	усвоен хорошо	усвоен отлично
БИК-16	17	4	6	5	2
БИК-17	19	4	8	4	3
БИК-18	21	1	5	9	6



Рис. 5. Сравнение результатов промежуточной успеваемости в контрольных группах по дисциплине «Теория сигналов»
Fig. 5. Comparison of the results of intermediate progress in the control groups in the discipline "Theory of Signals"

Основным критерием оценивания успешности применения игрофикации в учебном процессе стали оценки по результатам учебного семестра, где (из 100 баллов):

- материал не усвоен (оценка «2») – менее 61 балла;
- материал усвоен удовлетворительно (оценка «3»), – от 61 до 75 баллов включительно;
- материал усвоен хорошо («4») – от 76 до 90 баллов включительно;
- материал усвоен отлично («5») – от 91 до 100 баллов включительно.

Результаты промежуточной успеваемости в контрольных группах по дисциплине «Теория сигналов» приведены на рис. 5. Итогом проведения апробации внедрения игровых механик на примере дисциплины «Теория сигналов» стало значительное увеличение студентов с оценкой «отлично», не менее чем в два раза, а также рост общих показателей успеваемости до 83,4 балла (см. рис. 5). Приведенные показатели свидетельствуют о том, что внедрение игровых форм обучения прямо влияют на показатели успеваемости внутри группы.

Заключение

Подводя итог, отметим, что выделение оптимальных игровых механик для улучшения процесса обучения проходил по субъективным критериям, направленным на студентов технических специальностей. Таким образом, повторные исследования для групп иных специальностей может дать иные, но не менее интересные результаты.

Установлено, что применение игровых механик, и в целом игрофикации, при проведении лабораторного практикума, объективно повышает успеваемость, тем самым увеличивая сформированность умений и навыков в сфере профессиональных компетенций.

Список литературы

1. **Олесева М. М., Афанасьева С. Р.** Теоретические аспекты проблемы формирования познавательной активности студентов // Педагогический журнал. 2018. Т. 8, № 1А. С. 9–18.
2. **Олесева М. М., Афанасьева С. Р.** Активные и интерактивные формы обучения в высшей школе // III Ломоносовские чтения. Актуальные вопросы фундаментальных и прикладных исследований: Сб. ст. Междунар. науч.-практ. конф. Петрозаводск: Новая наука, 2019. С. 16–24.
3. **Левченкова Т. В., Черенкова И. А.** Активные методы как средства обеспечения интерактивной формы обучения в высшем образовании // Informatization of Higher education: current situation and development prospects: Materials of the III International scientific conference. Prague: Sociosféra-CZ, 2017. С. 17–21.
4. **Мазелис А. Л.** Геймификация в электронном обучении // Качество профессиональной подготовки через диалог системы высшего образования и профессиональной среды. Владивосток: ВГУЭС, 2013. 124 с.
5. **Белоус И. А., Чупалов А. Я.** Сравнительный анализ современных систем дистанционного обучения // Вестник Московского городского педагогического университета. Серия: Информатика и информатизация образования. 2019. № 3. С. 85–95.

References

1. **Olesova M. M., Afanasyeva S. R.** Theoretical aspects of the problem of the formation of students' cognitive activity. *Pedagogical Journal*, 2018, vol. 8, no. 1A, pp. 9–18. (in Russ.)
2. **Olesova M. M., Afanasyeva S. R.** Active and interactive forms of education in higher education. In: III Lomonosov readings. Topical issues of fundamental and applied research. Collection of articles of the International Scientific and Practical Conference. Petrozavodsk, New Science, 2019, pp. 16–24. (in Russ.)
3. **Levchenkova T. V., Cherenkova I. A.** Active methods as a means of providing an interactive form of education in higher education. In: Informatization of Higher education: current situation and development prospects: Materials of the III International scientific conference. Prague, Sociosféra-CZ, 2017, pp. 17–21. (in Russ.)
4. **Mazelis A. L.** Gamification in e-learning. In: Quality of vocational training through a dialogue between the higher education system and the professional environment. Vladivostok, VSUES, 2013, 124 p. (in Russ.)
5. **Belous I. A., Chupalov A. Ya.** Comparative analysis of modern distance learning systems. *Bulletin of the Moscow City Pedagogical University. Series: Informatics and informatization of education*, 2019, no. 3, pp. 85–95. (in Russ.)

Информация об авторах

Игорь Александрович Белоус, кандидат физико-математических наук, доцент
Михаил Валерьевич Ковырнёв, студент 3 года обучения
Михаил Владимирович Клюкман, студент 3 года обучения
Ирина Александровна Пяткова, студент 3 года обучения
Дмитрий Геннадьевич Сорока, студент 3 года обучения

Information about the Authors

Igor A. Belous, Candidate of Sciences (Physics and Mathematics), Associate Professor
Mikhail V. Kovyrnev, 3rd year student
Mikhail V. Klyukman, 3rd year student
Irina A. Piatkova, 3rd year student
Dmitriy G. Soroka, 3rd year student

*Статья поступила в редакцию 10.06.2021;
одобрена после рецензирования 01.08.2021; принята к публикации 01.08.2021
The article was submitted 10.06.2021;
approved after reviewing 01.08.2021; accepted for publication 01.08.2021*

Научная статья

УДК 004.8

DOI 10.25205/1818-7900-2021-19-3-26-39

Семантическая сегментация изображений в проекте «Duckietown»

Дарья Евгеньевна Шабалина¹
Кристина Сергеевна Ланчуковская²
Татьяна Викторовна Лях³
Константин Владимирович Чайка⁴

¹⁻³ Новосибирский государственный университет
Новосибирск, Россия

⁴ Санкт-Петербургский государственный электротехнический университет
Санкт-Петербург, Россия

⁴ Лаборатория алгоритмов мобильных роботов «JetBrains Research»
Санкт-Петербург, Россия

¹ d.shabalina@g.nsu.ru, <https://orcid.org/0000-0003-2418-6172>

² k.lanchukovskaya@g.nsu.ru, <https://orcid.org/0000-0002-9085-3634>

³ t.liakh@g.nsu.ru, <https://orcid.org/0000-0001-9148-946X>

⁴ pro100kot14@gmail.com, <https://orcid.org/0000-0001-5778-9266>

Аннотация

Статья посвящена сравнительному анализу алгоритмов семантической сегментации и исследованию их применимости на примере проекта «Duckietown». Рассмотрены как классические алгоритмы семантической сегментации, так и алгоритмы, использующие подходы машинного обучения. Исследованы фреймворки машинного обучения с учетом всех ограничений проекта «Duckietown». По результатам исследования для решения задачи сегментации в проекте «Duckietown» были выбраны нейросетевые алгоритмы, основанные на сетях U-Net, SegNet, DeepLab-v3, FC-DenseNet и PSPNet. U-Net и SegNet и протестированы на симуляторе «Duckietown».

Ключевые слова

робототехника, компьютерное зрение, нейронные сети, семантическая сегментация изображений, проект «Duckietown», роботы, искусственный интеллект

Благодарности

Работа выполнена при поддержке Лаборатории алгоритмов мобильных роботов «JetBrains Research»

Для цитирования

Шабалина Д. Е., Ланчуковская К. С., Лях Т. В., Чайка К. В. Семантическая сегментация изображений в проекте «Duckietown» // Вестник НГУ. Серия: Информационные технологии. 2021. Т. 19, № 3. С. 26–39. DOI 10.25205/1818-7900-2021-19-3-26-39

© Шабалина Д. Е., Ланчуковская К. С., Лях Т. В., Чайка К. В., 2021

ISSN 1818-7900 (Print). ISSN 2410-0420 (Online)

Вестник НГУ. Серия: Информационные технологии. 2021. Том 19, № 3. С. 26–39

Vestnik NSU. Series: Information Technologies, 2021, vol. 19, no. 3, pp. 26–39

Semantic Image Segmentation in Duckietown

Daria E. Shabalina¹, Kristina S. Lanchukovskaya²
Tatyana V. Lyakh³, Konstantin V. Chaika⁴

¹⁻³ Novosibirsk State University
Novosibirsk, Russian Federation

⁴ Saint Petersburg Electrotechnical University
St. Petersburg, Russian Federation

⁴ Mobile Robot Algorithms Laboratory “JetBrains Research”
St. Petersburg, Russian Federation

¹ d.shabalina@g.nsu.ru, <https://orcid.org/0000-0003-2418-6172>

² k.lanchukovskaya@g.nsu.ru, <https://orcid.org/0000-0002-9085-3634>

³ t.lyakh@g.nsu.ru, <https://orcid.org/0000-0001-9148-946X>

⁴ pro100kot14@gmail.com, <https://orcid.org/0000-0001-5778-9266>

Abstract

The article is devoted to evaluation of the applicability of existing semantic segmentation algorithms for the “Duckietown” simulator. The article explores classical semantic segmentation algorithms as well as ones based on neural networks. We also examined machine learning frameworks, taking into account all the limitations of the “Duckietown” simulator. According to the research results, we selected neural network algorithms based on U-Net, SegNet, DeepLab-v3, FC-DenseNet and PSPNet networks to solve the segmentation problem in the “Duckietown” project. U-Net and SegNet have been tested on the “Duckietown” simulator.

Keywords

robotics, computer vision, neural networks, semantic image segmentation, Duckietown, duckiebots, artificial intelligence

Acknowledgements

The work was carried out with the support of the Mobile Robot Algorithms Laboratory “JetBrains Research”

For citation

Shabalina D. E., Lanchukovskaya K. S., Liakh T. V., Chaika K. V. Semantic Image Segmentation in Duckietown. *Vestnik NSU. Series: Information Technologies*, 2021, vol. 19, no. 3, p. 26–39. (in Russ.) DOI 10.25205/1818-7900-2021-19-3-26-39

Введение

В настоящее время активно развиваются и используются тренажеры, которые имитируют различные роботизированные системы, а также промышленные технологические системы. Одним из таких тренажеров является «Duckietown»¹. Это открытый исследовательский проект в области автономно управляемых роботов. Он состоит из двух частей – роботов (duckiebots) и города (Duckietown), по которому они передвигаются. В городе есть размеченные дороги, светофоры, дорожные знаки и препятствия. Сами роботы – duckiebots – представляют собой двухколесную платформу, которая управляется двумя моторами постоянного тока, основным вычислительным модулем является одноплатный компьютер Raspberry Pi. На роботах установлены драйвер моторов и одна камера с Fisheye объективом². Несмотря на техническую простоту, это позволяет роботу двигаться согласно разметке, избегать столкновений с препятствиями и другими роботами, определять свое местоположение на карте, самостоятельно передвигаться по городу, взаимодействовать с другими роботами.

Тренажер отражает особенности массовых промышленных автопилотируемых систем с ограниченными ресурсами: такие роботы оснащены только камерами, а также имеют сла-

¹ URL: <https://www.duckietown.org/> (дата обращения 31.03.2021).

² Getting the Duckiebot hardware. URL: https://docs.duckietown.org/daffy/opmanual_duckiebot/out/get_db_hw.html (дата обращения 31.03.2021).

бую вычислительную мощность и минимальный набор датчиков. Подобные роботы используются в промышленности, например, в качестве роботов-погрузчиков [1].

На сегодняшний день важной задачей, в том числе в рамках проекта «Duckietown», является семантическая сегментация изображений, которые поступают с камеры робота. Причем данная задача применима как к данным камеры реального мира, так и к симуляционным. Семантическая сегментация [2] – это задача поиска групп пикселей изображения, каждая из которых характеризует один смысловой объект. При семантической сегментации требуется не только обнаружить и классифицировать объекты, но и определить их границы. Она применяется во многих сферах, в том числе и для алгоритмов управления беспилотным транспортом, а также в различных практических сферах, таких как промышленная робототехника, автомобильная промышленность, медицина [3]. В таких системах семантическая сегментация позволяет точнее проанализировать спутниковые изображения и карты, рентгенографические, электронно-микроскопические снимки, изображения, полученные с камер автопилотируемых автомобилей и роботов [4; 5]. Сложность решения этой задачи для тренажера «Duckietown» обусловлена ограничениями, которые накладывает робот: малой вычислительной мощностью, отсутствием дополнительных датчиков, монокулярной камерой.

Статья посвящена сравнительному анализу алгоритмов семантической сегментации и их применимости на примере проекта «Duckietown».

Обзор методов и алгоритмов семантической сегментации изображений

Для решения задач семантической сегментации применяются следующие подходы:

- граничные алгоритмы;
- пороговые алгоритмы;
- алгоритмы наращивания областей;
- алгоритмы сегментации на основе кластеризации;
- нейронные сети.

При анализе подходы оценивались по следующим критериям:

- устойчивость;
- оптимальные условия;
- ограничения и применение в робототехнике;
- ограничения и применение в автономном вождении.

Граничные алгоритмы

Граничные методы [6] основаны на выделении контуров объектов. Работа граничных методов состоит из двух этапов: поиск граничных пикселей с помощью операторов Робертса, Собеля, Лапласа, Кэнни и т. д., и замыкание границ. Основным недостатком граничных методов является сильная зависимость от «шума», что зачастую ведет к пересегментации [7; 8].

Пороговые алгоритмы

Пороговые алгоритмы [9] основаны на разделении гистограммы изображения на различные части по пороговым значениям. Существуют два различных метода: метод с глобальным порогом и метод с адаптивным порогом. Остальные алгоритмы – это производные от этих двух методов. Пороговые алгоритмы используют, когда нужно провести сегментацию простого изображения, области которого значительно различаются по яркости. Но при наличии различных помех – вспышек, теней, неравномерного освещения, что часто встречается в реальной жизни, пороговые методы не дают хорошего результата.

Алгоритмы наращивания областей

Алгоритмы наращивания областей в простейшем варианте основаны на выборе пикселя и изучении смежных с ним пикселей для проверки значений на близость по яркости [10]. Ес-

ли значения яркости близки, то соответствующие пиксели относятся в одну группу. Таким образом, в результате сращивания пикселей формируется область. Существуют и более эффективные варианты, в них в качестве начальной точки используются не отдельные пиксели, а небольшие области изображения. Каждая область проверяется на однородность: если результат проверки неудовлетворительный, то соответствующая область разбивается на более мелкие. Процесс продолжается до тех пор, пока все области не пройдут проверку на однородность. Далее производится формирование областей с помощью наращивания небольших областей, а не отдельных пикселей. Недостатком алгоритма является задача выбора начальной точки для сегментируемой области вручную, поэтому метод не может быть полностью автоматизирован.

Алгоритмы сегментации на основе кластеризации

Кластеризация [11] – задача разбиения множества объектов на непересекающиеся подмножества, называемые кластерами, таким образом, чтобы каждый кластер состоял из схожих объектов, а объекты разных кластеров существенно отличались. Входными данными для задачи кластеризации являются признаковые описания объектов – наборы характеристик, свойственные каждому объекту. Выходными данными служат кластеры (непересекающиеся множества). Чтобы привести задачу сегментации к задаче кластеризации, нужно отобразить точки изображения в пространство признаков, а также ввести в этом пространстве признаков метрику. К недостаткам кластеризации можно отнести то, что она плохо работает на зашумленных изображениях.

Нейронные сети

Нейронные сети [12] в упрощенном виде можно представлять как способ моделирования в технических системах принципов организации и механизмов функционирования головного мозга человека. В настоящее время наиболее популярным решением проблемы семантической сегментации является использование нейронных сетей. Благодаря им точность семантической сегментации значительно увеличилась. Среди основных характеристик ИНС выделяют обучение на примерах и обобщение полученных знаний, подавление шумов, отказоустойчивость.

Результаты сравнения алгоритмов семантической сегментации приведены в табл. 1.

Выводы по результатам проведенного исследования

Для решения задачи сегментации в Duckietown было решено использовать нейросетевые алгоритмы, так как они отличаются устойчивостью и быстродействием, а также уже использовались в робототехнике и автономном вождении. Для предобработки изображений было предложено использовать классические алгоритмы компьютерного зрения. Были отобраны алгоритмы наращивания областей и алгоритмы кластеризации. Пороговые алгоритмы были отклонены, поскольку успешность применения данного метода зависит от освещения. Граничные алгоритмы не подойдут как самостоятельные алгоритмы для разметки, так как требуется не только выделить границы, но и закрасить их, однако они отлично работают в паре с другими алгоритмами.

Обзор нейронных сетей

Современная тенденция – использование глубоких нейронных сетей для сегментации изображений. В обзоре сравнивались нейронные сети, предназначенные для семантической сегментации изображений, по точности, преимуществам, ограничениям и области применения. Результат сравнения приведен в табл. 2.

Таблица 1
 Алгоритмы семантической сегментации
 Table 1
 Semantic segmentation algorithms

Алгоритмы сегментации изображения	Устойчивость	Дополнительные ограничения	Оптимальные условия	Область применения
Граничные алгоритмы	Любой шум может повлиять на результат сегментации; низкий контраст изображения может влиять на качество обнаружения границ	Выявляет все границы, и бывает сложно найти корреляцию между интересующей областью и полученными границами	Изображения с высоким контрастом между различными областями	Используются в паре с другими алгоритмами сегментации
Пороговые алгоритмы	Успешность применения данного метода зависит от освещения и от правильно выбранного порога	Зависимость результатов сегментации от выбранной величины порога; повышенная чувствительность к шуму и неоднородностям в интенсивности	Простые изображения с высоким контрастом между различными областями	Чаще используется для выбора интересующей области изображения, игнорируя ненужные части; один из этапов в процедуре обработки МРТ-данных; робототехника
Алгоритмы наращивания областей	Нарращивание области может быть чувствительно к шуму	Точность и качество сегментации зависят от выбора начальных параметров; необходимость в выборе начальной точки для каждой сегментированной области производится вручную – следовательно, метод не может быть полностью автоматизирован	Если известно расположение стартовых точек, а также грамотно выбран критерий однородности, то результаты качественные	Робототехника
Алгоритмы на основе кластеризации	Могут быть чувствительны к шуму и неоднородностям в интенсивности	Зависит от конкретного алгоритма	Зависит от конкретного алгоритма	Сегментация спутниковых изображений, сегментация цветного изображения рака желудка
Нейронные сети	Обладают повышенной устойчивостью	Зависит от конкретной нейросети	Зависит от конкретной нейросети	Практически во всех видах сегментации

Таблица 2

Сравнение нейросетевых алгоритмов

Table 2

Compration of neural network algorithms

Сеть для сегментации	Точность, mIoU	Достоинства	Ограничения	Область применения
FCN-8	~ 57	Хорошая производительность; быстрое действие	Необходимость адаптации для сегментации; невысокая точность сегментации	Спутниковые изображения
U-Net	~ 77.5	Хорошие результаты сегментации даже на небольших датасетах; высокая точность; быстрая	Невысокая точность сегментации для некоторых задач	Duckietown; медицина; дорожные сцены
SegNet	~ 56.1	Высокая производительность; высокая скорость работы	Необходимость адаптации для сегментации; низкая точность сегментации	Duckietown; классификация типов дорожных знаков; дорожные сцены
FC-DenceNet	~ 66.9	Хорошая точность	Недостаточная точность сегментации для некоторых задач	Дорожные сцены
RefineNet	~ 73.6	Высокая точность сегментации	Невысокое быстродействие	Применяется везде, в том числе и для сегментации дорог
PSPNet	~ 81.2			Дорожные сцены; медицина
DeepLab-v3	~ 81.3			Спутниковые изображения
LinkNet	~ 76.4			
ESPNet	~ 60.3	Работает в реальном времени	Необходим мощный процессор	
ERFNet	~ 68.0			
ICNet	~ 69.5			
ENet	~ 58.3			Дорожные сцены

1. *FCN* [13] реализует архитектуру «кодер – декодер». В качестве кодера используется CNN без полносвязных слоев, выполняющая операции свертки и пулинга, декодер же использует транспонированные свертки. На входе сети – изображение произвольного размера, на выходе – карта сегментации.

Достоинствами данной сети являются хорошая производительность и ее быстродействие. Но FCN имеет некоторые существенные ограничения, а именно: необходимость адаптировать нейронную сеть для сегментации, низкое разрешение карт сегментации на выходе и невысокая точность сегментации по современным меркам. Использовалась для семантической сегментации спутниковых изображений.

2. *U-Net* [14] реализует архитектуру «кодер – декодер», базируясь на FCN. В случае U-Net кодер, как и у предыдущей сети, состоит из слоев свертки и пулинга, но в отличие от FCN у U-Net на каждом этапе удваивается количество каналов карт признаков. Декодер же состоит из слоев, повышающих разрешение карт признаков и слоев свертки, которые уменьшают количество каналов признаков.

Достоинствами данной сети являются хорошие результаты сегментации даже на небольших датасетах, высокое качество карт сегментации на выходе, быстрота. Но U-Net имеет недостаточную точность сегментации для некоторых задач, например при определении дорожных знаков. Сфера применения U-Net достаточно обширная и включает медицину, дорожные сцены. Также эта сеть уже использовалась в Duckietown.

3. *SegNet* [15; 16] – SegNet, как и U-Net, реализует архитектуру «кодер – декодер». Кодер состоит из слоев свертки и пулинга, а декодер – из слоев анпулинга и свертки. В отличие от других сетей, в этой сети применяются более быстрые и эффективные скуп-соединения.

Достоинствами данной сети являются высокая производительность и скорость работы, а также небольшое потребление оперативной памяти компьютера. Но SegNet характеризуют некоторые ограничения, например необходимость адаптировать сеть для сегментации, недостаточно хорошее разрешение карт сегментации и низкая точность. Применялась для классификации типов дорожных знаков и для сегментации дороги. Также эта сеть уже использовалась в Duckietown.

4. *DeepLab* [17] реализует архитектуру CNN вместе с многомасштабной обработкой изображений для получения карт сегментации. Для увеличения количества захватываемой информации без увеличения числа параметров в DeepLab используются разряженные ядра свертки.

Достоинствами данной сети являются высокое качество карт сегментации на выходе и высокая точность сегментации. Но DeepLab имеет невысокое быстродействие. Применялась для сегментации дорог, а также в медицине для сегментирования опухолей головного мозга.

5. *FC-DenseNet*³ расширяет архитектуру U-net, реализованную на архитектуре «кодер – декодер». Кодер сети состоит из классификационной сверточной нейронной сети DenseNet без полносвязных слоев. В отличие от других сетей, в FC-DenseNet реализованы DenseBlock, а также используются длинные скуп-соединения между кодером и декодером и короткие – между отдельными этапами кодера.

Достоинствами данной сети являются высокое качество карт сегментации на выходе и хорошая точность сегментации. При этом FC-DenseNet имеет недостаточную точность сегментации для некоторых задач. Применяется для сегментации дорожных сцен.

6. *RefineNet*⁴ включает в себя идеи полностью сверточных нейронных сетей и многомасштабной обработки изображений. Кодером является часть классификационной сверточной

³ Review: FC-DenseNet – One Hundred Layers Tiramisu, Fully Convolutional DenseNet. URL: <https://towardsdatascience.com/review-fc-densenet-one-hundred-layer-tiramisu-semantic-segmentation-22ee3be434d5> (дата обращения 31.03.2021).

⁴ Review: RefineNet – Multi-path Refinement Network. URL: <https://towardsdatascience.com/review-refinenet-multi-path-refinement-network-semantic-segmentation-5763d9da47c1> (дата обращения 31.03.2021).

нейронной сети ResNet. Декодер базируется на специальных блоках, предназначенных для обработки и объединения карт высокого разрешения кодера и карт низкого разрешения из предыдущего блока декодера.

Достоинствами данной сети являются высокое качество карт сегментации на выходе и высокая точность сегментации. Но RefineNet имеет невысокое быстродействие. Применялась везде, в том числе и для сегментации дорожных сцен.

7. *PSPNet* [18] базируется на архитектуре «кодер – декодер» и многомасштабной обработке с использованием пирамиды изображений. Кодер сети основывается на нейронной сети ResNet, в которую дополнительно входят разреженные свертки. Декодер сети использует модуль пулинга пирамиды.

Достоинствами данной сети являются высокое качество карт сегментации на выходе и высокая точность сегментации. Но PSPNet имеет невысокое быстродействие. Применялась в медицинских целях, например для сегментации изображений опухолей головного мозга, а также применялась для сегментации дорог.

8. *ENet* [19] реализует архитектуру «кодер – декодер». В качестве кодера используется сверточная нейронная сеть ResNet.

Достоинством данной сети является возможность работы в «реальном» времени. Главное ограничение ENet – необходимость иметь мощный процессор. Эта сеть разрабатывалась как раз для распознавания городских улиц.

9. *ICNet* [20] – каскадная нейронная сеть, основанная, как и RefineNet, на идеях многомасштабной обработки изображения. Архитектура ICNet базируется на топологии сети PSPNet.

Данная сеть имеет возможность работы в реальном времени. Главное ограничение ICNet – необходимость иметь мощный процессор. Используется для сегментации дорожных сцен.

10. *LinkNet* [21; 22] основана на архитектуре «кодер – декодер». Между блоками кодера и декодера сети имеются соединения, которые суммируют соответствующие карты признаков, агрегируя информацию с текущего и предыдущих этапов работы сети.

Достоинством данной сети является возможность работы в реальном времени. Главное ограничение LinkNet – необходимость иметь мощный процессор. Использовалась для семантической сегментации спутниковых изображений

11. *ESPNet* [23], как и сеть PSPNet, реализует архитектуру, включающую в себя идеи как архитектуры «кодер – декодер», так и многомасштабной обработки изображения в виде построения пирамиды карт признаков.

Данная сеть имеет возможность работы в реальном времени. Главное ограничение ESPNet – необходимость иметь мощный процессор. Используется для сегментации дорожных сцен.

12. *ERFNet* [24] объединяет идеи нескольких высокопроизводительных глубоких нейронных сетей, как Inception, Xception и ResNet, а также архитектуры «кодер – декодер» для решения задачи семантической сегментации на устройствах с ограниченными вычислительными ресурсами.

Достоинством данной сети является возможность работы в реальном времени. Главное ограничение ESPNet – необходимость иметь мощный процессор. Используется для сегментации дорожных сцен.

Исходя из данных табл. 2 и накладываемых ограничений (область применения, приемлемое отношение скорости работы к точности карт сегментации, малая мощность процессора Raspberry Pi), можно сделать следующие выводы:

- FCN-8, LinkNet не подходят, так как область применения не включает ничего, связанного с дорожным движением;
- RefineNet не подходит из-за невысокого быстродействия и относительно не самой высокой точности сегментации по отношению к сетям с таким же ограничением;
- для ENet, ICNet, ESPNet, ERFNet нужен мощный процессор.

В итоге самыми подходящими стали сети U-Net, SegNet, DeepLab-v3, FC-DenceNet [25] и PSPNet [26], которые подходят по всем критериям.

Применение нейронных сетей на примере Duckietown

Ранее уже предпринимались попытки использовать нейронные сети для семантической сегментации в тренажере «Duckietown», но в обоих случаях сегментация захватывала только дорожную разметку, игнорируя дорожные знаки, других duckiebots и “жителей” города.

- U-Net

В работе «Semantic segmentation transfer in Duckietown»⁵ обучающие данные были получены из симулятора Duckietown. Датасет собран вручную. Для данной задачи в нейронной сети U-Net уменьшено количество нейронных слоев. Затем модель была обучена на датасете, размеченном с помощью Photoshop.

В качестве оптимизатора использовался стохастический градиентный спуск (SGD) с импульсом (momentum) 0,9, скоростью обучения 0,01 и уменьшением веса 10⁻⁴.

На смоделированных изображениях точность проверки пикселей к концу обучения составила 98,42 %. Тестирование модели проходило как на смоделированных, так и на реальных изображениях.

- SegNet

В работе [28] задача сегментации изображений на тренажере решалась в 2018 г. с помощью нейронной сети SegNet и GAN⁶.

Обзор фреймворков

Для семантической сегментации с помощью нейросетевых алгоритмов были исследованы библиотеки машинного обучения (табл. 3). Во всех представленных фреймворках используется язык программирования Python. Каждая библиотека является opensource проектом.

Таблица 3

Сравнение фреймворков для нейросетевых алгоритмов

Table 3

Comparison of frameworks for neural network algorithms

Фреймворк	Требования	Область применения
TensorFlow	ОС: Linux, Mac OS X, Microsoft Windows, iOS, Android	Для проектов машинного обучения. Для проектов нейронных сетей. В автоматизированном программном обеспечении для создания титров, таких как DeepDream
Keras	ОС: Linux, Mac OS X, Microsoft Windows	Для исследования глубокого обучения. Популярный среди стартапов, разрабатывающих продукты, основанные на глубоком обучении
PyTorch	ОС: Linux Mac OS X Microsoft Windows, iOS, Android	Используется для решения различных задач: компьютерное зрение, обработка естественного языка

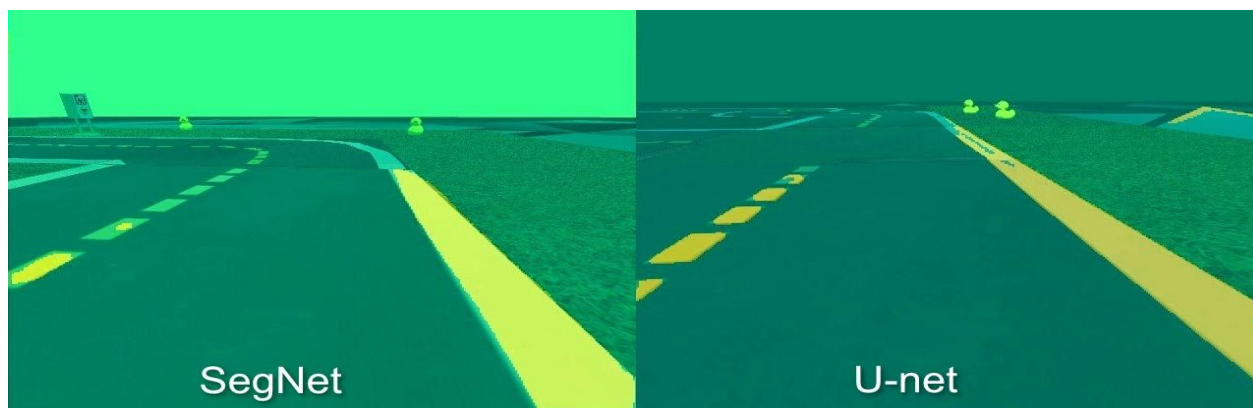
⁵ URL: <https://github.com/duckietown/segmentation-transfer> (дата обращения 31.03.2021).

⁶ Duckietown Simulator. URL: https://docs.duckietown.org/daffy/AIDO/draft/dt_simulator.html (дата обращения 31.03.2021).

В итоге были выбраны библиотеки машинного обучения Keras и PyTorch, так как порог вхождения у них намного ниже, чем у TensorFlow, а функции они выполняют те же самые.

Тестирование нейросетей SegNet и U-Net на изображениях с тренажера Duckietown

В эксперименте были получены изображения с тренажера Duckietown⁷ и сформирован датасет для обучения SegNet и U-Net. Целью эксперимента было сегментировать исходное изображение нейросетью на 2 класса – дорогу и фон. При схожих затратах по производительности, нейронная сеть U-net показала лучший результат:



Заключение

Проведен анализ применимости алгоритмов семантической сегментации в тренажере Duckietown, который имитирует автопилотируемых роботов в условиях городской среды. Были проанализированы подходы, основанные на классических алгоритмах компьютерного зрения, и подходы, основанные на нейросетевых алгоритмах. Установлено, что методы, основанные на классических алгоритмах компьютерного зрения, уступают методам, основанных на нейронных сетях, по устойчивости, точности сегментации и скорости работы. Было предложено использовать классические алгоритмы компьютерного зрения для разметки изображений и подготовки датасетов.

Для сегментации изображений на duckiebots было предложено использовать нейросети. Проведен анализ нейросетей, использующихся для сегментации, с учетом особенностей тренажера Duckietown (малая вычислительная мощность, отсутствие дополнительных датчиков, монокулярная камера).

Таким образом, для сегментации изображений были выбраны нейросетевые алгоритмы, основанные на сетях U-Net, SegNet, DeepLab-v3, FC-DenceNet и PSPNet. Для предобработки изображений выбраны классические алгоритмы компьютерного зрения, такие как алгоритмы наращивания областей и алгоритмы кластеризации. Для апробации на тренажере Duckietown были выбраны библиотеки машинного обучения Keras и PyTorch.

Лучший результата в ходе апробации получен с использованием U-net.

В дальнейшем планируется апробация DeepLab-v3, FC-DenceNet и PSPNet, а также подготовка разработанного решения сегментации изображений к развертыванию на реальном роботе.

⁷ Duckietown Simulator. URL: https://docs.duckietown.org/daffy/AIDO/draft/dt_simulator.html (дата обращения 31.03.2021).

Список литературы

1. **Sharma Y.** Adoption of next generation robotics: A case study on Amazon. Pune, 2017, 15 p.
2. **Csurka G., Larlus D., Perronnin F.** What is a good evaluation measure for semantic segmentation. Meylan, 2013, 11 p.
3. **Зуенко Е. А., Шулунова А. А.** Автоматическое выделение боковых желудочков головного мозга на МРТ изображениях // Вестник НГУ. Серия: Информационные технологии. 2017. Т. 15, № 4. С. 22–31. DOI 10.25205/1818-7900-2017-15-4-22-31
4. **Мусатян С. А., Ломакин А. В., Саргасов С. Ю., Попыванов Л. К., Монахов И. Б., Чиждова А. С.** Способы сегментации медицинских изображений // Тр. Ин-та системного программирования РАН. М., 2018. DOI 10.15514/ISPRAS-2018-30(4)-12
5. **Горбачева В. А., Криворотова И. А., Маркелов А. О., Котлярова Е. В.** Семантическая сегментация спутниковых снимков аэропортов с помощью свёрточных нейронных сетей. Мю, 2019. 10 с.
6. **Ozmen A., Akman E. T.** Edge detection using steerable filters and CNN. In: 11th European Signal Processing Conference. Toulouse, 2002, pp. 1–4.
7. **Pestunov I., Rylov S., Berikov V.** Hierarchical Ensemble Clustering Algorithm for Multispectral Image Segmentation. In: Proceedings 9th Open German-Russian Workshop on Pattern Recognition and Image Understanding (OGRW-2014). Koblenz, Uni. of Koblenz-Landau Press, 2015, pp. 123–127.
8. **Рылов С. А., Мельников П. В., Пестунов И. А.** Спектрально-текстурная классификация гиперспектральных изображений высокого пространственного разрешения // Интерэкспо ГЕО-Сибирь. 2016. Т. 4, № 1. С. 78–84.
9. **Чичварин Н. В.** Пороговые методы. М., 2016.
10. **Луценко М. Т., Ульянычев Н. В., Семичевская Н. П.** Автоматизированная система распознавания объектов на микроскопических изображениях биологических образцов. Благовещенск, 1999.
11. **Пестунов И. А., Синявский Ю. Н.** Алгоритмы кластеризации в задачах сегментации спутниковых изображений // Вестник КемГУ. 2012. Т. 52, № 4/2. С. 110–125.
12. **Li B., Shi Y., Qi Z., Chen Z.** A Survey on Semantic Segmentation. Beijing, 2018. DOI 10.1109/ICDMW.2018.00176
13. **Noh H., Hong S., Han B.** Learning deconvolution network for semantic segmentation. Pohang, 2015.
14. **Ronneberger O., Fischer P., Brox T.** U-Net: Convolutional Networks for Biomedical Image Segmentation. Berlin, 2015.
15. **Badrinarayanan V., Kendall A., Cipolla R.** SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, vol. 39, no. 12, pp. 2481–2495. DOI 10.1109/TPAMI.2016.2644615
16. **Емельянов А. В.** Анализ методов семантической сегментации изображений на основе нейронных сетей. М., 2019.
17. **Chen L., Papandreou G., Kokkinos I., Murphy K., Yuille A. L.** DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, vol. 40, no. 4, pp. 834–848. DOI 10.1109/TPAMI.2017.2699184
18. **Zhou J., Hao M., Zhang D., Zou P., Zhang W.** Fusion PSPnet Image Segmentation Based Method for Multi-Focus Image Fusion. *IEEE Photonics Journal*, 2019, vol. 11, no. 6, pp. 1–12. DOI 10.1109/JPHOT.2019.2950949
19. **Paszke A.** ENet: A Deep Neural Network Architecture for Real-Time Semantic Segmentation. Ithaca, 2016, 10 p.
20. **Zhao H., Qi X., Shen X., Shi J., Jia J.** ICNet for Real-Time Semantic Segmentation on High-Resolution Images. Hong Kong, 2018, 16 p.

21. **Chaurasia A., Culurciello E.** LinkNet: Exploiting encoder representations for efficient semantic segmentation. *EEE Visual Communications and Image Processing (VCIP)*, 2017, pp. 1–4. DOI 10.1109/VCIP.2017.8305148
22. **Емельянов А. В., Галиахметов Д. Г.** Анализ методов семантической сегментации изображений в реальном времени на основе нейронных сетей. М., 2020.
23. **Mehta S., Rastegari M., Caspi A., Shapiro L., Hajishirzi H.** ESPNet: Efficient Spatial Pyramid of Dilated Convolutions for Semantic Segmentation. Seattle, 2018, 29 p.
24. **Romera E., Alvarez J. M., Bergasa L. M., Arroyo R.** ERFNet: Efficient Residual Factorized ConvNet for Real-time Semantic Segmentation. *IEEE Transactions on Intelligent Transportation Systems*, 2018, vol. 19, no. 1, pp. 263–272. DOI 10.1109/TITS.2017.2750080
25. **Brahimi S., Aoun N. B., Chokri B. A., Benoit A., Lambert P.** Multiscale Fully Convolutional DenseNet for Semantic Segmentation. Pilsen, 2018.
26. **Li X.** Examining the spatial distribution and temporal change of the green view index in New York City using Google Street View images and deep learning. Philadelphia, 2020, 16 p.
27. **Tóth Z.** Deep Learning-based Semantic Segmentation in Simulation and Real-World for Autonomous Vehicles. URL: <https://smartlabai.medium.com/deep-learning-based-semantic-segmentation-in-simulation-and-real-world-for-autonomous-vehicles-e7fe25cef816> (accessed 31.03.2021).
28. **Айрапетов А. Э., Коваленко А. А.** Исследование генеративно-состязательной сети // Политехнический молодежный журнал МГТУ им. Н. Э. Баумана. 2018. № 10. DOI 10.18698/2541-8009-2018-10-380

References

1. **Sharma Y.** Adoption of next generation robotics: A case study on Amazon. Pune, 2017, 15 p.
2. **Csurka G., Larlus D., Perronnin F.** What is a good evaluation measure for semantic segmentation. Meylan, 2013, 11 p.
3. **Zuenko E. A., Shulunova A. A.** Automated Segmentation of the Lateral Ventricles from MRI Image. *Vestnik NSU. Series: Information Technologies*, 2017, vol. 15, no. 4, pp. 22–31. (in Russ.) DOI 10.25205/1818-7900-2017-15-4-22-31
4. **Musatyan S. A., Lomakin A. V., Sartasov S. Yu., Popyvanov L. K., Monakhov I. B., Chizhova A. S.** Methods for segmentation of medical images. In: Proceedings of the Institute for System Programming of the Russian Academy of Sciences. Moscow, 2018. (in Russ.) DOI 10.15514/ISPRAS-2018-30(4)-12
5. **Gorbacheva V. A., Krivorotova I. A., Markelov A. O., Kotlyarova E. V.** Semantic segmentation of satellite images of airports using convolutional neural networks. Moscow, 2019. 10 p. (in Russ.)
6. **Ozmen A., Akman E. T.** Edge detection using steerable filters and CNN. In: 11th European Signal Processing Conference. Toulouse, 2002, pp. 1–4.
7. **Pestunov I., Rylov S., Berikov V.** Hierarchical Ensemble Clustering Algorithm for Multispectral Image Segmentation. In: Proceedings 9th Open German-Russian Workshop on Pattern Recognition and Image Understanding (OGRW-2014). Koblenz, Uni. of Koblenz-Landau Press, 2015, pp. 123–127.
8. **Rylov S. A., Melnikov P. V., Pestunov I. A.** Spectral and texture classification of high spatial resolution hyperspectral images. *Interexpo GEO-Siberia*, 2016, vol. 4, no. 1, pp. 78–84. (in Russ.)
9. **Chichvarin N. V.** Threshold methods. Moscow, 2016. (in Russ.)
10. **Lutsenko M. T., Ulyanychev N. V., Semichevskaya N. P.** Automatic system for object recognition in microscopic images of biological samples. Blagoveshchensk, 1999. (in Russ.)
11. **Pestunov I. A., Sinyavsky Yu. N.** Clustering algorithms in satellite images segmentation tasks. *Vestnik KSU*, 2012, vol. 52, no. 4/2, pp. 110–125. (in Russ.)

12. **Li B., Shi Y., Qi Z., Chen Z.** A Survey on Semantic Segmentation. Beijing, 2018. DOI 10.1109/ICDMW.2018.00176
13. **Noh H., Hong S., Han B.** Learning deconvolution network for semantic segmentation. Pohang, 2015.
14. **Ronneberger O., Fischer P., Brox T.** U-Net: Convolutional Networks for Biomedical Image Segmentation. Berlin, 2015.
15. **Badrinarayanan V., Kendall A., Cipolla R.** SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, vol. 39, no. 12, pp. 2481–2495. DOI 10.1109/TPAMI.2016.2644615
16. **Emelyanov A. V.** Analysis of Semantic Image Segmentation Methods Based on Neural Networks. Moscow, 2019. (in Russ.)
17. **Chen L., Papandreou G., Kokkinos I., Murphy K., Yuille A. L.** DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, vol. 40, no. 4, pp. 834–848. DOI 10.1109/TPAMI.2017.2699184
18. **Zhou J., Hao M., Zhang D., Zou P., Zhang W.** Fusion PSPnet Image Segmentation Based Method for Multi-Focus Image Fusion. *IEEE Photonics Journal*, 2019, vol. 11, no. 6, pp. 1–12. DOI 10.1109/JPHOT.2019.2950949
19. **Paszke A.** ENet: A Deep Neural Network Architecture for Real-Time Semantic Segmentation. Ithaca, 2016, 10 p.
20. **Zhao H., Qi X., Shen X., Shi J., Jia J.** ICNet for Real-Time Semantic Segmentation on High-Resolution Images. Hong Kong, 2018, 16 p.
21. **Chaurasia A., Culurciello E.** LinkNet: Exploiting encoder representations for efficient semantic segmentation. *EEE Visual Communications and Image Processing (VCIP)*, 2017, pp. 1–4. DOI 10.1109/VCIP.2017.8305148
22. **Emelyanov A. V., Galiahmetov D. G.** Analysis of Semantic Image Segmentation Methods in Real Time Based on Neural Networks. Moscow, 2020. (in Russ.)
23. **Mehta S., Rastegari M., Caspi A., Shapiro L., Hajishirzi H.** ESPNet: Efficient Spatial Pyramid of Dilated Convolutions for Semantic Segmentation. Seattle, 2018, 29 p.
24. **Romera E., Alvarez J. M., Bergasa L. M., Arroyo R.** ERFNet: Efficient Residual Factorized ConvNet for Real-time Semantic Segmentation. *IEEE Transactions on Intelligent Transportation Systems*, 2018, vol. 19, no. 1, pp. 263–272. DOI 10.1109/TITS.2017.2750080
25. **Brahimi S., Aoun N. B., Chokri B. A., Benoit A., Lambert P.** Multiscale Fully Convolutional DenseNet for Semantic Segmentation. Pilsen, 2018.
26. **Li X.** Examining the spatial distribution and temporal change of the green view index in New York City using Google Street View images and deep learning. Philadelphia, 2020, 16 p.
27. **Tóth Z.** Deep Learning-based Semantic Segmentation in Simulation and Real-World for Autonomous Vehicles. URL: <https://smartlabai.medium.com/deep-learning-based-semantic-segmentation-in-simulation-and-real-world-for-autonomous-vehicles-e7fe25cef816> (accessed 31.03.2021).
28. **Ajrabetov A. E., Kovalenko A. A.** Research of the generative adversarial network. *Polytechnic Youth Journal of the Moscow State Technical University. N. E. Bauman*, 2018, no. 10. (in Russ.) DOI 10.18698/2541-8009-2018-10-380

Информация об авторах

Дарья Евгеньевна Шабалина, студент бакалавриата, 3 курс
Кристина Сергеевна Ланчуковская, студент бакалавриата, 3 курс
Татьяна Викторовна Лях, кандидат технических наук
Константин Владимирович Чайка, аспирант 3 года обучения

Information about the Authors

Daria E. Shabalina, Bachelor Student
Kristina S. Lanchukovskaya, Bachelor Student
Tatyana V. Lyakh, Candidate of Sciences (Engineering)
Konstantin V. Chaika, Post-Graduate Student

*Статья поступила в редакцию 20.06.2021;
одобрена после рецензирования 01.08.2021; принята к публикации 01.08.2021
The article was submitted 20.06.2021;
approved after reviewing 01.08.2021; accepted for publication 01.08.2021*

Научная статья

УДК 004.550

DOI 10.25205/1818-7900-2021-19-3-40-49

Облачный сервис многомерной обработки количественных данных для решения геологических задач

Алексей Андреевич Загумёнов¹
Вера Викторовна Наумова²
Виталий Сергеевич Ерёмченко³

¹ Институт автоматизации и процессов управления
Дальневосточного отделения Российской академии наук
Владивосток, Россия

^{2,3} Государственный геологический музей им. В. И. Вернадского
Российской академии наук
Москва, Россия

¹ truepikvic@gmail.com, <https://orcid.org/0000-0002-0501-5362>

² v.naumova@sgm.ru, <https://orcid.org/0000-0002-3001-1638>

³ vitaer@gmail.com, <https://orcid.org/0000-0002-5250-5743>

Аннотация

В работе описан разработанный облачный веб-сервис многомерной обработки количественных данных для решения широкого класса научных геологических задач. Вычислительный узел «Многомерные методы анализа данных» позволяет выполнять обработку табличных данных различными методами современного анализа данных с настройкой их параметров и визуализацией результатов. Узел включает в себя такие группы методов, как предобработка данных, описательная статистика, кластерный анализ, факторный анализ, корреляционный анализ, регрессионный анализ и др. Вычислительный узел «Многомерные методы анализа данных» входит в состав Вычислительно-аналитической геологической среды ГГМ РАН и интегрирован с ее сервисами. При этом вычислительный узел является самостоятельным облачным веб-сервисом, взаимодействие с которым происходит посредством REST API. Это позволяет обращаться к многомерным методам анализа данных, размещенным на вычислительном узле, широкому кругу пользователей, в том числе интегрировать его в другие информационные системы как стороннее приложение для обработки табличных данных.

Ключевые слова

облачные сервисы, геология, вычислительный узел, REST API

Благодарности

Работа выполняется в рамках Государственного задания ГГМ РАН по теме № 0140-2019-0005 «Разработка информационной среды интеграции данных естественнонаучных музеев и сервисов их обработки для наук о Земле»

Для цитирования

Загумёнов А. А., Наумова В. В., Ерёмченко В. С. Облачный сервис многомерной обработки количественных данных для решения геологических задач // Вестник НГУ. Серия: Информационные технологии. 2021. Т. 19, № 3. С. 40–49. DOI 10.25205/1818-7900-2021-19-3-40-49

© Загумёнов А. А., Наумова В. В., Ерёмченко В. С., 2021

ISSN 1818-7900 (Print). ISSN 2410-0420 (Online)

Вестник НГУ. Серия: Информационные технологии. 2021. Том 19, № 3. С. 40–49

Vestnik NSU. Series: Information Technologies, 2021, vol. 19, no. 3, pp. 40–49

Cloud Service for Multidimensional Processing of Quantitative Data for Solving Geological

Aleksei A. Zagumennov¹, Vera V. Naumova²
Vitaliy S. Eremenko³

¹ Institute of Automation and Control Processes
of the Far East Branch of the Russian Academy of Sciences

^{2,3} Vladivostok, Russian Federation

² V. I. Vernadsky State Geological Museum of the Russian Academy of Sciences
Moscow, Russian Federation

¹ truepikvic@gmail.com, <https://orcid.org/0000-0002-0501-5362>

² v.naumova@sgm.ru, <https://orcid.org/0000-0002-3001-1638>

³ vitaer@gmail.com, <https://orcid.org/0000-0002-5250-5743>

Abstract

The study describes the developed cloud web service for multidimensional processing of quantitative data for solving a wide class of scientific geological tasks. The computing node “Multidimensional methods of data analysis” provides processing of tabular data using various methods of modern data analysis and allows to set their parameters and visualize the results. The node includes wide range of methods such as data preprocessing, descriptive statistics, cluster analysis, factor analysis, correlation analysis, regression analysis. Computing node “Multidimensional methods of data analysis” is a part of Computational analytical geological environment of State Geological Museum of RAS and is integrated with its services. At the same time, the computing node is an independent cloud web service which implements REST API for interaction with it. This allows a wide range of users to access multidimensional data analysis methods located on a computing node and provides capabilities of its integration into information systems as a third-party application for processing tabular data.

Keywords

cloud services, geology, computing node, REST API

Acknowledgements

The study is supported by the Government contract no. 0140-2019-0005 “Development of an information environment for integrating data from natural science museums and services for their processing for Earth sciences”

For citation

Zagumennov A. A., Naumova V. V., Eremenko V. S. Cloud Service for Multidimensional Processing of Quantitative Data for Solving Geological. *Vestnik NSU. Series: Information Technologies*, 2021, vol. 19, no. 3, p. 40–49. (in Russ.) DOI 10.25205/1818-7900-2021-19-3-40-49

Введение

При геологических исследованиях быстрыми темпами накапливается большое количество геологической информации: результаты геологической документации буровых скважин, горных выработок и естественных обнажений, спектральных и химических анализов руд, горных пород и минералов, данные геофизических и геохимических измерений и др. Одно из важнейших направлений научно-технического прогресса в геологии состоит в широком внедрении автоматизированных методов накопления, хранения, обработки и передачи геологической информации с целью повышения эффективности геологических исследований.

Сегодня для решения целого ряда геологических задач используются многомерные методы анализа данных. Данные методы применяются с целью обнаружения скрытых закономерностей, выявления наиболее существенных связей между переменными путем анализа взаимосвязи большого количества этих переменных. К таким методам относят факторный, кластерный, дисперсионный, регрессионный анализ, многомерное шкалирование, анализ канонических переменных и др.

На рынке представлено немало традиционных профессиональных пакетов, обеспечивающих эти функции для анализа геологических данных: SPSS, STATISTICA, MATLAB и др.

Среди современных веб-инструментов анализа данных (в том числе и Big Data) необходимо выделить нижеследующие.

Orange – проект, содержащий набор инструментов для визуализации, обработки и анализа данных. Данный пакет инструментов поможет создать интерактивный рабочий процесс для анализа и визуализации данных с помощью точечных диаграмм, гистограмм, деревьев, дендрограмм, сетей и тепловых карт.

NodeXL – программное обеспечение для анализа данных, визуализации зависимостей и сетей. Вот некоторые его возможности:

- визуализация собственной сети в соответствии с требованиями;
- анализ социальных сетей и определение ключевых подписчиков;
- анализ контента позволяет глубоко анализировать все слова, хэштеги и URL-адреса, встроенные в сообщения, твиты и т. д.;
- импорт наборов данных из различных источников;
- автоматизация и отчеты.

Salesforce reports & dashboards – профессиональный инструмент для анализа и визуализации больших наборов данных. Вот некоторые его особенности:

- фильтрация, суммирование, создание формул;
- визуализация данных с обновлением в режиме реального времени;
- управление данными из административной части CRM-системы.

Стоит отметить, что Salesforce не относится к бесплатным инструментам.

RapidMiner – еще один из современных инструментов. Комплекс включает в себя IDE собственной разработки и сервер для хранения выполненных процессов анализа, общего доступа к информации аналитиками и подготовки результатов для выдачи в веб-интерфейсе. Платформа позволяет быстро оперировать данными из разных источников, запускать проверки моделей машинного обучения, и всё это в одном месте с одним набором программного обеспечения.

Однако при решении геологических задач не всегда требуется полная функциональность профессиональных пакетов, а необходимы лишь отдельные методы обработки данных.

Основная цель работ – обеспечение геологов облачным сервисом, в единой точке доступа предоставляющим методы многомерной обработки количественных данных для решения геологических задач на основе открытых решений, для более эффективного использования, анализа распределенных геологических ресурсов и возможности построения прогнозов на основе геологических моделей.

Для этого был разработан Вычислительный узел «Многомерные методы анализа данных» в составе Вычислительно-аналитической геологической среды ГГМ РАН (далее Вычислительная среда).

Вычислительно-аналитическая геологическая среда ГГМ РАН

С 2017 г. в Государственном геологическом музее им В. И. Вернадского РАН ведется разработка и адаптация методов и технологий обработки и анализа территориально распределенной разнотипной геологической информации и сервисов ее обработки. На основе созданных подходов, разработанных методов и технологий реализуется информационно-аналитическая среда для поддержки и сопровождения научных исследований в геологии, осуществляющая интеграцию территориально распределенной геологической информации с использованием специализированных служб ее анализа и обработки (<http://geologyscience.ru>) (рис. 1) [1]. Авторы предполагают, что разработанная программная платформа управления тематическими сервисами обработки и анализа, которая является частью информационно-аналитической среды, обеспечит пользователям доступ к хранилищам современных наукоемких алгоритмов и вычислительных ресурсов, необходимых для оперативной обработки больших массивов геологических данных.

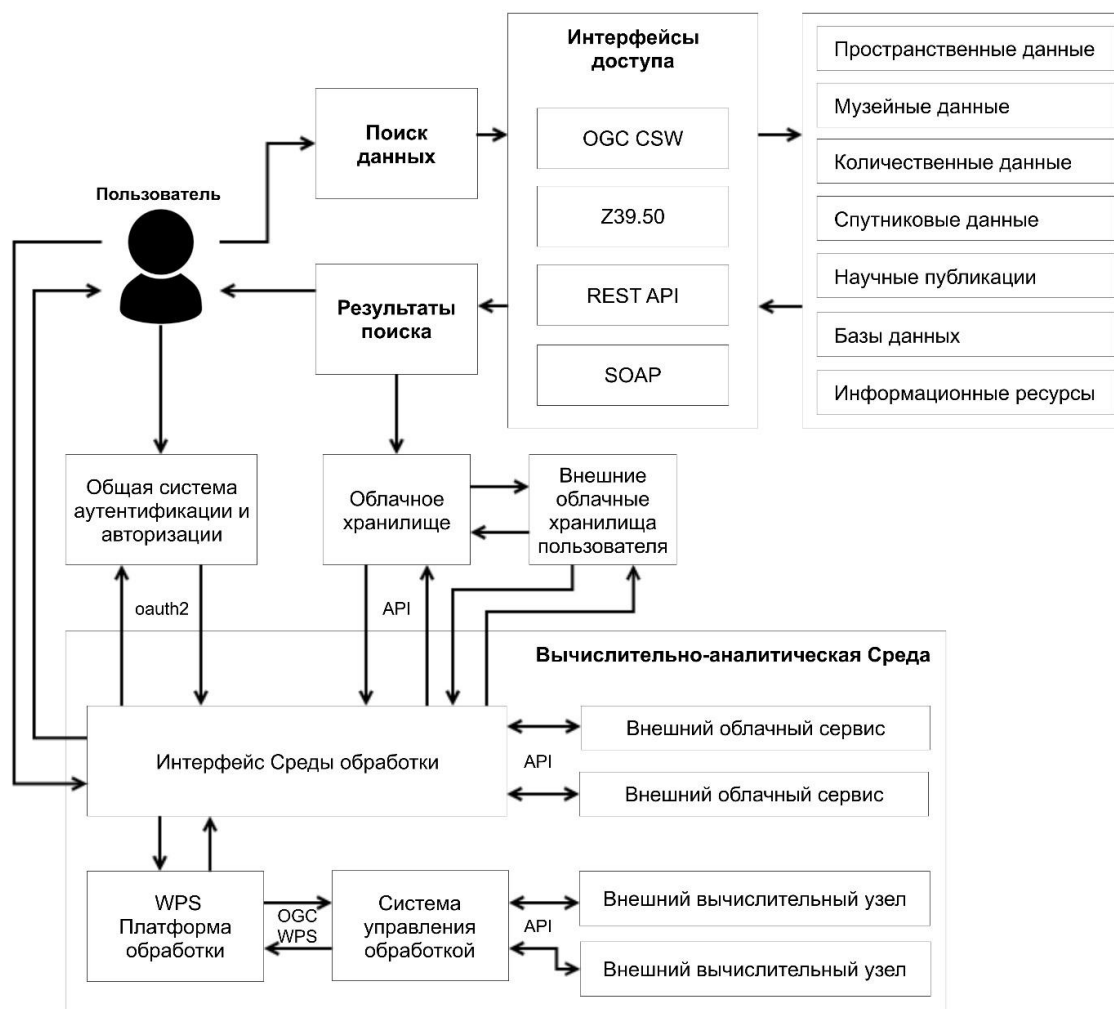


Рис. 1. Общая функциональная схема Информационно-аналитической среды GeologyScience.ru

Fig. 1. General functional diagram of the Informational-analytical environment GeologyScience.ru

В Вычислительной среде реализован доступ к различным вычислительным узлам и платформам [2; 3].

При разработке Вычислительной среды разработаны подходы к организации единого рабочего пространства исследователя для обработки геологической информации (<https://service.geologyscience.ru>). Предложены подходы по взаимодействию с внешними территориально распределенными сервисами обработки и анализа геологических данных на основе WPS платформы и к организации интегральной платформы для доступа к интерактивным облачным сервисам обработки геологических данных.

В результате проведенных работ авторами разработана цифровая платформа по обработке и анализу геологических данных DPAP [3]. Платформа представляет собой программный продукт на языке программирования Java в формате веб-приложения с использованием фреймворка Spring Boot. Она реализована с использованием принципов слабосвязанной архитектуры. Для взаимодействия с пользователем реализован веб-интерфейс с JavaScript фреймворком Angularjs. При этом разработан программный интерфейс с использованием архитектурного стиля REST для взаимодействия с платформой в формате «система – система».

Серверная технология DPAR обеспечивает безопасное подключение к Вычислительной среде совместной работы, которая превращает пользовательские разрозненные данные в безопасный для анализа фонд информации. Масштабируемая технология создания каталогов и управления большими наборами данных позволяет администрировать одновременный доступ с разными пользовательскими настройками и допусками безопасности. Платформа позволяет пользователям эффективно каталогизировать, управлять, визуализировать большие пространственные данные, что поможет им, минимизировать риск и принимать более обоснованные решения, увеличивающие возможность успеха обнаружения месторождений.

Использование цифровой платформы по обработке и анализу геологических данных позволяет в дальнейшем расширять возможности разработчиков информационно-аналитических сред для поддержки и сопровождения научных исследований.

В Вычислительной среде реализован доступ к следующим вычислительным узлам и платформам.

- **Вычислительный узел «Многомерные методы анализа данных»**, разрабатываемый в ГИМ РАН, который позволяет выполнять обработку табличных данных различными методами анализа данных с настройкой их параметров и визуализацией результатов. Вычислительный узел включает в себя такие группы методов, как предобработка данных, описательная статистика, кластерный, факторный, корреляционный, регрессионный анализ и др.

- **Обработка петролого-геохимических данных.** В Институте физики Земли РАН разработана интерактивная база методов обработки петролого-геохимических данных [4]. Данная система предоставляет сервисы построения спайдерграмм, гистограмм и классификационных диаграмм; сервис идентификации минералов по их химическому составу; сервис интерпретации состава минерала и разложение на минералы и т. д. Интерфейс взаимодействия с сервисами построен на основе REST архитектуры.

- **Структурный анализ публикаций.** В междисциплинарном центре математического и вычислительного моделирования (Университет Варшавы, Польша) разработан сервис для извлечения метаданных из научных публикаций [5]. Метаданные включают в себя авторов, аффилиацию, аннотацию, ключевые слова, название журнала, объем, год выпуска, разобранные библиографические ссылки, структуру разделов документа, заголовки разделов и абзацы. Интерфейс взаимодействия с сервисами построен на основе REST архитектуры.

- **Обработка естественного языка.** В Университете Шеффилда в рамках проекта GATE (General Architecture for Text Engineering) разработан ряд сервисов по обработке текстовых данных для различных языков [6]. Для обработки текстовых данных на русском языке предоставляются сервисы по определению частей речи слов, а также выделению именованных сущностей, таких как имена и фамилии, названия организаций, географические названия, даты, денежные единицы и т. д. Интерфейс взаимодействия с сервисами построен на основе REST архитектуры.

В Вычислительную среду также включены следующие облачные сервисы. Одним из наиболее популярных программных продуктов обработки табличных данных в интерактивном режиме является **Excel** из пакета Microsoft Office. Данный программный продукт содержит ряд инструментов для редактирования данных, построения различных диаграмм, использования встроенных процедур анализа и создания собственных. Компания «Microsoft» разработала бесплатную облачную версию продукта Excel (<https://www.office.com/>). Зарегистрировавшись, можно использовать полноценную веб-версию для данных, расположенных в облачном хранилище Microsoft Onedrive.

Для анализа пространственных данных наиболее подходящим инструментом являются геоинформационные системы (ГИС), позволяющие в интерактивном режиме взаимодействовать с объектами на карте, применяя к отдельным слоям и объектам доступные средства анализа. Наиболее подходящим для геологии облачным решением является **ArcGIS Online**, разрабатываемый компанией ESRI. Portal for ArcGIS – это инфраструктура ArcGIS Online, функционирующая в защищенной ИТ-среде или в частном облаке организации (под контро-

лем сетевого экрана или в полностью изолированной локальной сети). Портал позволяет создавать карты, каталогизировать и анализировать пространственные данные с помощью удобного, интуитивного понятного интерфейса.

Для анализа спутниковых данных одним из лидеров среди облачных сервисов является **An Open Source Geospatial Data Management & Analysis Platform**. Open Data Cube (ODC) – это некоммерческий проект с открытым исходным кодом, который был создан для обеспечения доступа, управления и анализа больших объемов данных ГИС по мониторингу Земли. Он представляет собой общий аналитический фреймворк, содержащий наборы структурированных данных и инструментов, которые позволяют проводить анализ больших коллекций пространственных данных. ODC был разработан для анализа пространственных данных на больших временных промежутках, однако его можно использовать на любых наборах пространственных данных. Данные могут представлять собой модели высот, геофизические сетки, интерполированные поверхности и т. д. Ключевой особенностью ODC является возможность сохранения каждого уникального элемента набора пространственных данных, в отличие от многих других методов работы с большими коллекциями пространственных данных.

Вычислительный узел «Многомерные методы анализа данных»

Вычислительный узел позволяет выполнять обработку табличных данных различными методами анализа данных с настройкой их параметров и визуализацией результатов. Вычислительный узел включает в себя такие группы методов, как предобработка данных, описательная статистика, кластерный, факторный, корреляционный, регрессионный анализ и др. (рис. 2) (<https://service.geologyscience.ru/service?nodeId=1>). Перечень предлагаемых пользователю методов постоянно пополняется.

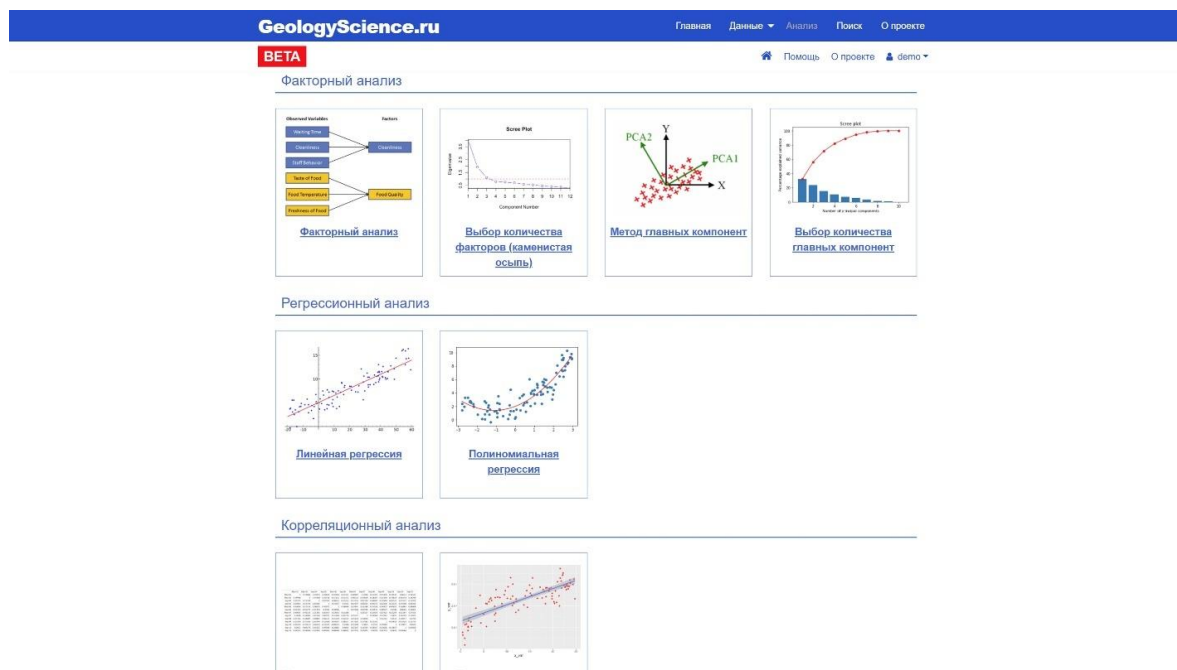


Рис.2. Страница выбора многомерных методов анализа данных

Fig. 2. Multidimensional methods of data analysis selection webpage

Выбор конкретных методов для включения в вычислительный узел обуславливается ориентированностью узла в первую очередь на решение геологических задач. Работа с геологическими данными часто опирается на статистические методы анализа данных, проверку гипотез, поиск зависимостей и корреляций. Результаты применения этих методов позволяют сделать выводы об исследуемых данных и принять решения о применении других методов: кластеризации, классификации и др. При этом важно иметь возможность настройки параметров методов для более детального анализа геологических данных.

Вычислительный узел состоит из следующих компонент (рис. 3):

- веб-сервис для получения запросов от внешних пользователей или систем на обработку данных выбранным методом с указанными параметрами, проверку статуса обработки и получения результатов работы метода;
- реестр методов, содержащий перечень методов, их параметров, а также непосредственную реализацию методов в виде скриптов на языке Python;
- очередь задач, в которую помещаются поступающие запросы на обработку данных, имеющую возможность отслеживания статуса выполнения задач;
- исполнители задач, которые при получении новой задачи из очереди обращаются к реестру методов и выполняют соответствующий скрипт;
- хранилище временных файлов, в котором для каждой задачи хранятся входные файлы и файлы результатов.

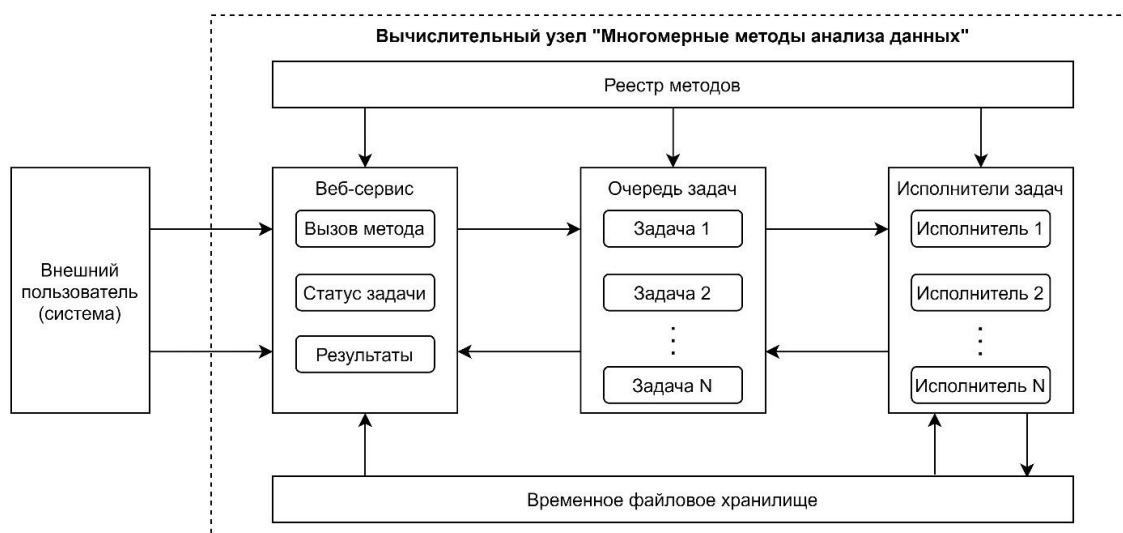


Рис. 3. Принципиальная схема Вычислительного узла «Многомерные методы анализа данных»

Fig. 3. Schematic diagram of the Computing node "Multidimensional methods of data analysis"

Вычисления производятся в среде Python с использованием известных открытых пакетов обработки данных: Scikit-learn, Pandas, Matplotlib и др. Обработка поступающих запросов на обработку данных производится фреймворком FastAPI посредством REST API, реализованного по стандарту OpenAPI (<http://spec.openapis.org/oas/v3.0.3>), с использованием очереди вычислительных задач, реализованной на основе NoSQL базы данных Redis.

При этом отдельные компоненты системы (веб-сервис, база данных с очередью задач, исполнители задач) разворачиваются с применением контейнеров Docker. Такая архитектура позволяет разделить обработку запросов и тяжелые вычисления больших объемов данных, что обеспечивает отказоустойчивость и масштабируемость узла.

Вычислительный узел «Многомерные методы анализа данных» является самостоятельным облачным веб-сервисом (<https://analysis.geologyscience.ru>), что позволяет использовать его отдельно, посредством REST API, реализованного по стандарту OpenAPI 3.0. Это обеспечивает возможность обращаться к многомерным методам анализа данных, размещенным на вычислительном узле, широкому кругу пользователей, в том числе интегрировать его в другие информационные системы как стороннее приложение для обработки табличных данных. Кроме того, использование стандарта OpenAPI делает узел самодокументированным, а при изменении списка методов документация обновляется автоматически (<https://analysis.geologyscience.ru/docs>). Согласно классификации облачных сервисов, приведенных в работе [7], вычислительный узел «Многомерные методы анализа данных» реализует модель «программное обеспечение как сервис» (Software as a Service, SaaS), т. е. пользователи сервиса получают доступ к возможностям сервиса посредством облачных технологий, без необходимости устанавливать и конфигурировать программное обеспечение самостоятельно.

Взаимодействие с Вычислительно-аналитической средой по обработке и анализу геологических данных

Взаимодействие с платформой по обработке и анализу геологических данных осуществляется с использованием разработанного REST интерфейса. В Вычислительной среде реализован модуль взаимодействия с внешними веб-сервисами (WSIM), предоставляющий единый интерфейс доступа к внешним сервисам обработки и анализа, работающим по принципу «запрос – ответ». В качестве интерфейса доступа используется OGC Web Processing Service 1.0. Каждый используемый Вычислительной средой метод публикуется в виде отдельного процесса внутри WPS сервиса. Для реализации WPS сервиса используется программный пакет с открытым исходным кодом GeoServer. GeoServer содержит встроенные компоненты для реализации очереди заданий, разграничения доступа к различным процессам и возможность как синхронного, так и асинхронного выполнения процессов. Для реализации модуля используются программные библиотеки GeoTools и 52 North WPS. Взаимодействие платформы с данным модулем осуществляется через WPS интерфейс.

В Вычислительной среде реализован доступ к общему облачному хранилищу пользователя, используемому в информационно-аналитической системе GeologyScience.ru, при помощи отдельного программного модуля (MICS). Хранилище позволяет размещать пользовательские данные для их дальнейшего использования в других подсистемах, в том числе для обработки и анализа. Результаты обработки также могут быть сохранены пользователем в облачное хранилище на длительный срок для их последующего использования в других сервисах. Взаимодействие с платформой осуществляется посредством REST интерфейса.

Для проведения предварительного анализа количественных данных в Вычислительной среде используется программный продукт Excel с доступом через интерактивную облачную платформу Microsoft Office Online, описанный ранее.

Все используемые в Вычислительной среде сервисы, включая сервисы вычислительного узла «Многомерные методы анализа данных», публикуются в каталоге сервисов среды. Основные функции каталога: обеспечение пользователя информацией об имеющихся в Вычислительной среде сервисах с возможностью получения информации о каждом зарегистрированном сервисе, включая общее описание, данных о поставщике сервиса и технической информации для организации взаимодействия. Информация из данного модуля используется в модуле мониторинга для отслеживания состояния каждого внешнего сервиса, зарегистрированного в Вычислительной среде. Взаимодействие с платформой осуществляется с использованием REST интерфейса.

Дальнейшие перспективы развития Вычислительного узла

Дальнейшее развитие Вычислительного узла предполагается вести в двух контекстах: прикладном и системном. В рамках прикладного контекста планируется расширение количества предлагаемых методов анализа данных, добавление пространственных методов обработки геологических данных, реализация механизма цепочек обработки геологических данных. Развитие узла в системном контексте предполагает реализацию возможности двусторонней работы с узлом обработки (по механизму Webhook или протоколу WebSocket), создание средств авторизации и учета пользователей узла, разработку модуля аналитики поведения пользователей, расширение возможностей интеграции вычислительного узла в другие информационные системы.

Список литературы

1. **Наумова В. В., Платонов К. А., Еременко В. С., Патук М. И., Дьяков С. Е.** Информационно-аналитическая среда для поддержки научных исследований в геологии: текущее состояние и перспективы развития // Тр. XVII Междунар. конф. «Распределенные информационно-вычислительные ресурсы. Цифровые двойники и большие данные. (DICR-2019)». Новосибирск, 2019. С. 139–147. DOI 10.25743/ICT.2019.70.61.021
2. **Eremenko V. S., Naumova V. V., Platonov K. A., Dyakov S. E., Eremenko A. S.** The main components of a distributed computational and analytical environment for the scientific study of geological systems. *Russian Journal of Earth Sciences*, 2018, vol. 18, iss. 6. DOI 10.2205/2018ES000636
3. **Еременко В. С., Наумова В. В., Загумёнов А. А., Булов С. В.** Облачные технологии для развития территориально распределенной вычислительно-аналитической геологической среды // Вычислительные технологии. 2021. Т. 26, № 1. С. 86–98.
4. **Иванов С. Д.** Интерактивный реестр геосенсоров на основе веб-приложения // Компьютерные исследования и моделирование. 2016. Т. 8, № 4. С. 621–632.
5. **Tkaczyk D., Szostek P., Fedoryszak M., Dendek P., Bolikowski L.** CERMINE: automatic extraction of structured metadata from scientific literature. *International Journal on Document Analysis and Recognition*, 2015, vol. 18, no. 4, pp. 317–335.
6. **Maynard D., Bontcheva K., Augenstein I.** Natural Language Processing for the Semantic Web. *Synthesis Lectures on the Semantic Web: Theory and Technology*, 2016, vol. 6, no. 2, pp. 1–194.
7. **Moser L., Thuraisingham B., Zhang J.** Services in the cloud. *IEEE transactions on services computing*, 2015, vol. 8, no. 2, pp. 172–174.

References

1. **Naumova V. V., Platonov K. A., Eremenko V. S., Patuk M. I., Dyakov S. E.** Information and analytical environment for supporting scientific research in geology: current state and development prospects. In: Proceedings of the XVII International conference “Distributed information and computing resources. (DICR-2019)”. Novosibirsk, 2019, pp. 139–147. (in Russ.) DOI 10.25743/ICT.2019.70.61.021
2. **Eremenko V. S., Naumova V. V., Platonov K. A., Dyakov S. E., Eremenko A. S.** The main components of a distributed computational and analytical environment for the scientific study of geological systems. *Russian Journal of Earth Sciences*, 2018, vol. 18, iss. 6. DOI 10.2205/2018ES000636
3. **Eremenko V. S., Naumova V. V., Zagumennov A. A., Bulov S. V.** Cloud technologies for development of geographically distributed computational and analytical Geological environment. *Computational Technologies*, 2021, vol. 26, no. 1, pp. 86–98. (in Russ.)

4. **Ivanov S. D.** Interactive Web Application Based Geosensors Registry. *Computer research and modeling*, 2016, vol. 8, no. 4, pp. 621–632. (in Russ.)
5. **Tkaczyk D., Szostek P., Fedoryszak M., Dendek P., Bolikowski L.** CERMINE: automatic extraction of structured metadata from scientific literature. *International Journal on Document Analysis and Recognition*, 2015, vol. 18, no. 4, pp. 317–335.
6. **Maynard D., Bontcheva K., Augenstein I.** Natural Language Processing for the Semantic Web. *Synthesis Lectures on the Semantic Web: Theory and Technology*, 2016, vol. 6, no. 2, pp. 1–194.
7. **Moser L., Thuraisingham B., Zhang J.** Services in the cloud. *IEEE transactions on services computing*, 2015, vol. 8, no. 2, pp. 172–174.

Информация об авторах

Алексей Андреевич Загумённый, научный сотрудник

Researcher ID R-4407-2016

Вера Викторовна Наумова, доктор геолого-минералогических наук

Researcher ID J-9039-2018

Виталий Сергеевич Ерёмченко, младший научный сотрудник

Researcher ID Q-3678-2016

Information about the Authors

Aleksei A. Zagumennov, Researcher

Researcher ID R-4407-2016

Vera V. Naumova, Doctor of Sciences (Geology and Mineralogy)

Researcher ID J-9039-2018

Vitaliy S. Eremenko, Junior Researcher

Researcher ID Q-3678-2016

Статья поступила в редакцию 20.06.2021;

одобрена после рецензирования 01.08.2021; принята к публикации 01.08.2021

The article was submitted 20.06.2021;

approved after reviewing 01.08.2021; accepted for publication 01.08.2021

Научная статья

УДК 004.932.2

DOI 10.25205/1818-7900-2021-19-3-50-60

Новая модель нейрона конца линий, применимая в свёрточных нейронных сетях

Александр Владимирович Кугаевских

Институт автоматизации и электрометрии
Сибирского отделения Российской академии наук
Новосибирск, Россия

Новосибирский государственный университет
Новосибирск, Россия

a-kugaevskikh@yandex.ru, <https://orcid.org/0000-0002-6676-0518>

Аннотация

Представлена модель комплексного нейрона конца линий. Этот тип нейронов дает максимальный отклик на конце линии и используется для уточнения выделения краев. В работе приведен обзор различных моделей нейронов конца линий и проведено сравнение их откликов. Нами предложена более простая и, в тоже время, точная *СЕ*-модель нейрона конца линий, базирующаяся на использовании фильтров Габора, находящихся в противофазе. Предложенная модель интегрирована в свёрточную нейронную сеть выделения краев, описание которой также приведено в статье. Предлагаемая модель показала высокую эффективность.

Ключевые слова

нейрон конца линий, выделение краев, нейронные сети, фильтр Габора, простые клетки, комплексные клетки

Благодарности

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 18-37-00029

Для цитирования

Кугаевских А. В. Новая модель нейрона конца линий, применимая в свёрточных нейронных сетях // Вестник НГУ. Серия: Информационные технологии. 2021. Т. 19, № 3. С. 50–60. DOI 10.25205/1818-7900-2021-19-3-50-60

New End-Stopped Complex Cell Model Applicable in Convolutional Neural Networks

Alexander V. Kugaevskikh

Institute of Automation and Electrometry
of the Siberian Branch of the Russian Academy of Sciences
Novosibirsk, Russian Federation

Novosibirsk State University
Novosibirsk, Russian Federation

a-kugaevskikh@yandex.ru, <https://orcid.org/0000-0002-6676-0518>

Abstract

This article is dedicated to modeling the end-stopped neuron. This type of neuron gives the maximum response at the end of the line and is used to refine the edge. The article provides an overview of different models of end-stopped neurons. I have proposed a simpler and more accurate model of an end-stopped neuron based on the use of Gabor filters

© Кугаевских А. В., 2021

ISSN 1818-7900 (Print). ISSN 2410-0420 (Online)

Вестник НГУ. Серия: Информационные технологии. 2021. Том 19, № 3. С. 50–60

Vestnik NSU. Series: Information Technologies, 2021, vol. 19, no. 3, pp. 50–60

in antiphase. For this purpose, the models of simple and complex cells whose output is used in the proposed model are also described. Simple cells are based on the use of a Gabor filter, the parameters of which are also described in this article. The proposed model has shown its effectiveness.

Keywords

end-stopped neuron, edge detection, neural network, Gabor filter, simple cell, complex cell

Acknowledgements

This paper was financially supported by the Russian Foundation for Basic Research (grant no. 18-37-00029)

For citation

Kugaevskikh A. V. New End-Stopped Complex Cell Model Applicable in Convolutional Neural Networks. *Vestnik NSU. Series: Information Technologies*, 2021, vol. 19, no. 3, p. 50–60. (in Russ.) DOI 10.25205/1818-7900-2021-19-3-50-60

Введение

При построении нейронных сетей, выполняющих функцию сегментации изображений, возникает задача точного определения границ. Для этого чаще всего используют фильтр Габор (в частности модели HMAX [1] и LPREEN [2] используют именно фильтр Габора). При этом возникает проблема выделения кривых. Уменьшая градус отклонения между ориентациями выделяемых линий, мы получаем повышение ошибки выделения из-за небольшого размера ядра фильтра Габора.

В качестве корректирующего механизма могут быть применены нейроны конца линий. Для зрительной коры нейроны конца линий (end-stopped complex cell) были описаны Хьюбелем и Вайзелем [3]. Особенность нейронов конца линий заключается в уменьшении активации нейрона при продолжении линии в пределах рецептивного поля (рис. 1). При этом помимо длины линии такие нейроны чувствительны к изменению контраста [4]. Позже выяснилось, что нейроны конца линий также участвуют в анализе движения [5], предоставляя дополнительную информацию для определения точного направления движения.

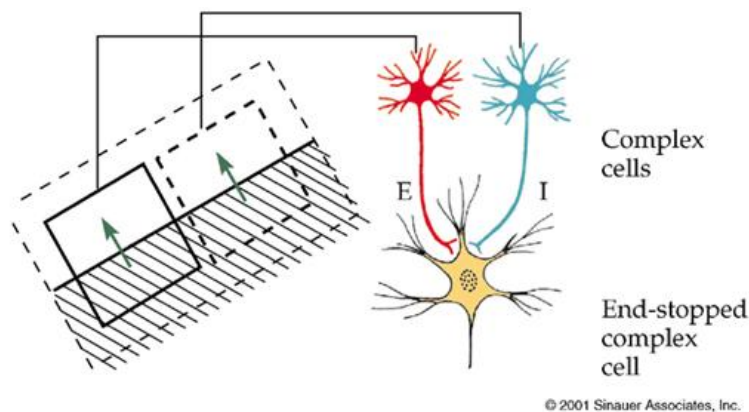


Рис. 1. Рецептивное поле нейрона конца линий
(E – возбуждающий нейрон, I – тормозящий нейрон) [6]
Fig. 1. Receptive field of the neuron of the end of the lines
(E – excitatory neuron, I – inhibitory neuron) [6]

Моделирование нейронов конца линий в нейронных сетях не только позволит повысить точность выделения границ, в том числе кривых, но и даст дополнительную информацию для выделения движения и распознавания иллюзорного контура. В некотором смысле нейроны конца линий выполняют ту же функцию, что и детекторы ключевых точек.

1. Выделение краев

Для выделения краев мы используем двухслойную нейронную сеть (рис. 2), предложенную нами в работе [7]. Выходы первого слоя такой сети помимо своей основной функции (выделения краев) служат также основой для работы нейронов конца линий. На первом слое выделяются линии определенной ориентации. Второй слой отвечает за выделение комбинаций линий, в том числе и углов. При этом связи между слоями организованы особым образом. Каждый нейрон второго слоя (U_{C2}) соединен только с двумя нейронами первого слоя (U_{S1}), рис. 3. Таким образом, нейроны второго слоя позволяют выделять линии и углы.

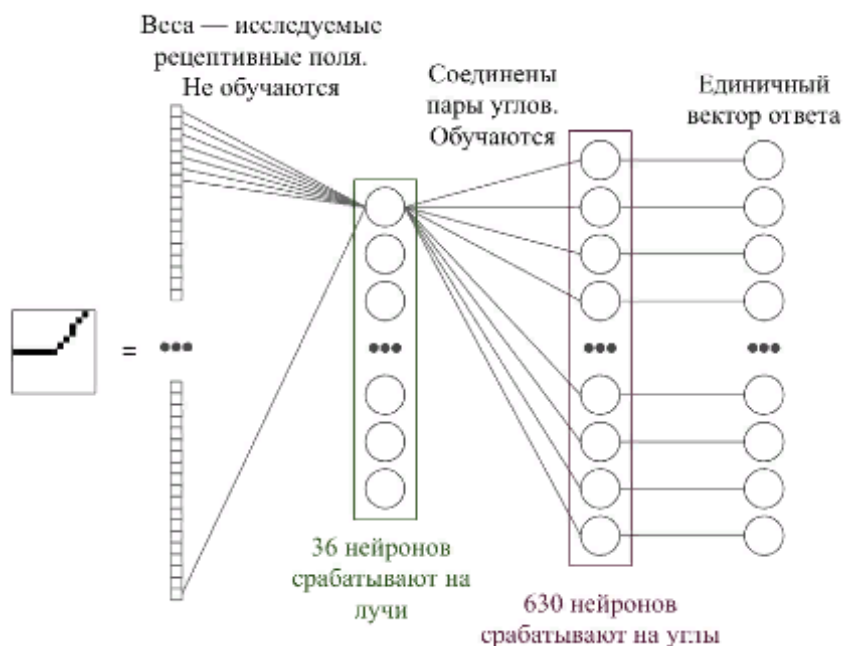


Рис. 2. Схема слоев выделения краев

Fig. 2. Scheme of edge selection layers

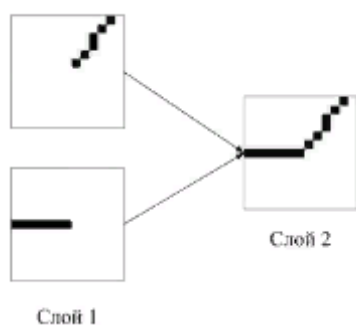


Рис. 3. Принцип организации связей между слоями

Fig. 3. The principle of the organization of links between layers

Нейроны первого слоя реагируют на линии предпочтительной ориентации θ . Их рецептивные поля формируются с помощью фильтра Габора. Фильтр Габора был выбран по двум причинам: во-первых, его функция позволяет формировать рецептивное поле простого нейрона для выделения линий определенной ориентации; во-вторых, в отличие от вейвлета «мексиканская шляпа» его постоянная компонента гораздо ближе к 0, что уменьшает паразитную активацию нейронов первого слоя. Равномерная заливка рецептивного поля такого

нейрона дает очень близкое к нулю значение. Фильтр Габора позволяет выделять светлые линии на темном фоне и темные линии на светлом фоне.

Изначально преобразование Габора предлагалось для фильтрации одномерного сигнала [8]. Даугман заметил сходство работы фильтра Габора и «простых» клеток первичной зрительной коры головного мозга [9].

Ядро фильтра Габора представляет собой произведения элемента базиса Фурье и гауссиана. За счет элемента базиса Фурье фильтр становится чувствителен к компонентам изображения определенной ориентации и пространственной частоты. Гауссиан необходим для пространственной локализации фильтра. В этом отношении свертка с фильтром Габора напоминает оконное преобразование Фурье, где функцией окна является гауссиан.

Формула Даугмана имеет специфический вид и не очень удобна для вычислений:

$$G = \exp(-\pi[(x - x_0)^2 a^2 + (y - y_0)^2 b^2]) \exp(-i2\pi(u(x - x_0) + v(y - y_0))), \quad (1)$$

где

(x_0, y_0) – позиция центра окна свертки,

(u, v) – вектор, определяющий предпочтительную пространственную частоту,

$a = \frac{1}{\sigma_x}, b = \frac{1}{\sigma_y}$ – параметры масштаба ядра фильтра.

Наложение фильтра на изображение осуществляется путем дискретной свертки с импульсной характеристикой или ядром фильтра с помощью сдвигающегося окна, поэтому для центра окна принимают значения $x_0 = 0, y_0 = 0$.

По формуле Эйлера экспоненту из формулы (1) можно переписать в виде

$$\exp(-i2\pi(ux + vy)) = \cos(2\pi(ux + vy)) + i\sin(2\pi(ux + vy)). \quad (2)$$

Такая запись фильтра образует квадратурную пару, состоящую из фильтров для определения симметричных и антисимметричных компонентов. При этом в большинстве случаев достаточно применения фильтра для определения симметричных компонент:

$$G_{\text{sym}}(x, y) = \exp\left(-\frac{x^2 + y^2}{\sigma^2}\right) \cos(2\pi(ux + vy)), \quad (3)$$

$$G_{\text{antisym}}(x, y) = \exp\left(-\frac{x^2 + y^2}{\sigma^2}\right) \sin(2\pi(ux + vy)). \quad (4)$$

Как правило, изображение фильтруют только по x -компоненте пространственной частоты, а остальное возлагается на поворот ядра фильтра. Таким образом, принимают $v = 0$. Также часто переходят от вектора (u, v) к фиксированной частоте $F = \sqrt{u^2 + v^2}$ (циклов / пиксель). У многих авторов вместо пространственной частоты фигурирует длина волны $\lambda = \frac{1}{F}$, определяющая размеры элементов изображения, на которые будет реагировать фильтр.

В литературе встречается множество вариаций формулы фильтра Габора с разными масштабирующими коэффициентами. Оптимальная формула фильтра Габора должна обладать низким значением постоянной компоненты (минимизация шума), и в то же время высокой селективной способностью (более четкое выделение краев на перепаде яркости, что дает более математически высокое значение свертки). Этим требованиям соответствует формула, предложенная Петковым [10; 11].

$$G_{1,2} = \exp\left(-\frac{x'^2 + y'^2}{2\sigma^2}\right) \cos\left(2\pi\frac{x'}{\lambda} + \phi\right), \quad (5)$$

где

ϕ – фаза, определяющая симметричность и чередование тормозных и возбуждающих областей ядра (в радианах),

σ – масштаб фильтра,

γ – степень эллиптичности фильтра.

Для реагирования на компоненты определенной ориентации в фильтр Габора вводятся параметры поворота координатной сетки на угол θ :

$$\begin{cases} x' = x \cos \theta + y \sin \theta, \\ y' = -x \sin \theta + y \cos \theta. \end{cases} \quad (6)$$

Так как фильтр имеет ярко выраженную ориентационную избирательность, градус отклонения между ориентациями ($\Delta\theta$) влияет на выявление деталей и, как следствие, уровень шума.

Параметр ϕ (см. (5)), отвечающий за симметричность ядра фильтра, был введен в качестве альтернативы квадратурной паре фильтров (см. (3) и (4)). Он может принимать значения $-\frac{\pi}{2}$ или $\frac{\pi}{2}$ для выявления антисимметричных компонент и $-\pi$, 0 или π для симметричных компонент; γ определяет вытянутость ядра фильтра по оси ординат. Параметр введен исключительно для упрощения подбора оптимального ядра без разных значений σ_x и σ_y и может быть выражен через эти параметры $\gamma = \frac{\sigma_x}{\sigma_y}$.

Многочисленная практика применения фильтра Габора доказала, что оптимальное значение σ подбирается из соотношения $\frac{\sigma}{\lambda}$, выведенного через пропускную способность. Впервые эта техника была предложена в [12].

$$\frac{\sigma}{\lambda} = \frac{1}{\pi} \sqrt{\frac{\ln 2}{2} \frac{2^B + 1}{2^B - 1}}. \quad (7)$$

По данным нейронауки, пропускная способность для нейронов первичной зрительной коры определяется в диапазоне $B = [0.6, 2.0]$ [13]. Зачастую принимают $B = 1$. В этом случае, $\frac{\sigma}{\lambda} \approx 0,56$.

Для задачи выделения краев фильтр Габора имеет следующие параметры $\gamma = 0.1$, $\theta \in [0, 350]$, $\lambda = 3$. В итоге, фильтр Габора позволяет выделять светлые линии на темном фоне ($\phi = -\pi$) и темные линии на светлом фоне ($\phi = 0$). Карта рецептивного поля нейрона на базе фильтра Габора представлена на рис. 4.

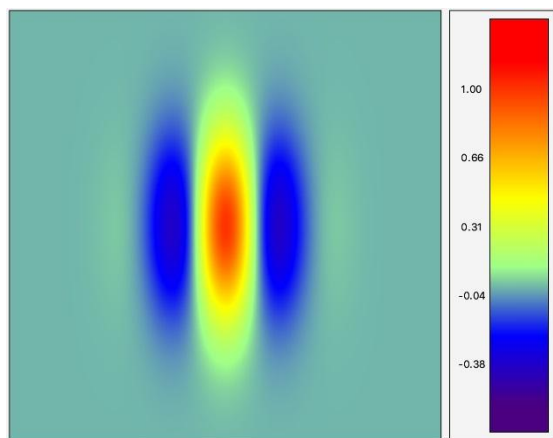


Рис. 4. Карта рецептивного поля нейрона на базе фильтра Габора

Fig. 4. Map of the receptive field of a neuron based on the Gabor filter

Нейроны первого слоя дают близкую по значению активацию для линий в пределах $\pm 20^\circ$ отклонения от предпочтительной ориентации. Это может создать проблему для точного определения ориентации признака на следующем слое.

Второй слой составляют сложные клетки, реагирующие на сочетания линий, углы и многоугольники. Хотя основная функция сложных клеток в зрительной коре – это реакция на движение, мы пока не моделируем данную функцию. Нейроны второго слоя получают входы от соответствующих по типу нейронов первого слоя. Эти связи являются обучаемыми. Обучение производилось с помощью обратного распространения ошибки. Для обучения и экспериментальной проверки поведения разных фильтров были сформированы пакеты черно-белых изображений. На рис. 5 представлена часть из 630 изображений для обучения нейрона выделения линий и углов.

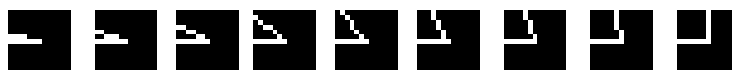


Рис. 5. Часть изображений для обучения и тестирования нейронов второго слоя на базе фильтра Габора

Fig. 5. Part of the images for training and testing neurons of the second layer based on the Gabor filter

На обучающей выборке было получено 100 % качество выделения сочетаний линий при условии минимального отклонения углов ориентации ($\Delta\theta = 10^\circ$) и размера рецептивного поля сложной клетки – 13 пикселей, таким образом, оптимальным размером ядра простой клетки на базе фильтра Габора является 7×7 пикселей. Такие параметры соотносятся с некоторыми нейрофизиологическими данными об устройстве нейронов первичной зрительной коры [14].

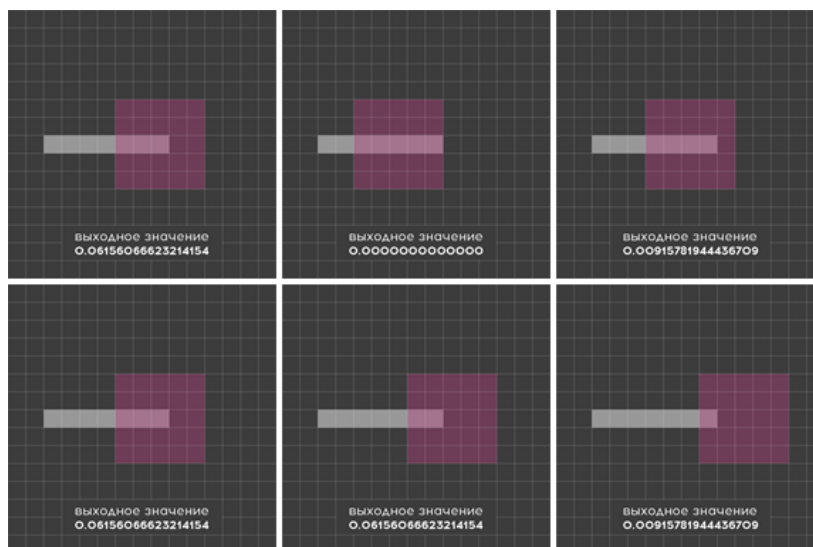
2. Обзор моделей нейронов конца линий

Нейроны конца линий моделировались разными способами. В B-V модели [15] (рис. 6) для выделения концов линий использовался комплексный вейвлет Морле:

$$ES1(x, y) = \frac{1}{4} x e^{-\left(\frac{x^2+y^2}{4} + \frac{k_1}{4}(k_1 - 2iy)\right)} \quad (8)$$

Рис. 6. Результат работы B-V модели

Fig. 6. The result of B-V model



Модель Родригеза и дю Баффа [16] основана на работах Вюртца [17] и Хейтгера [18] и использует разницу энергий Габора (рис. 7). Модель нейрона конца линий Хейтгера (10) использует энергию Габора (5), вычисляемую через симметричное и антисимметричное ядро фильтра Габора (3) и (4).

$$E(x, y) = \sqrt{G_{sym}^2(x, y) + G_{antisym}^2(x, y)}, \quad (9)$$

$$E_s(x, y) = f(E(x - d, y) - E(x + d, y)), \quad (10)$$

где $f(x) = \max(x, 0)$ – функция отсечения отрицательных значений.

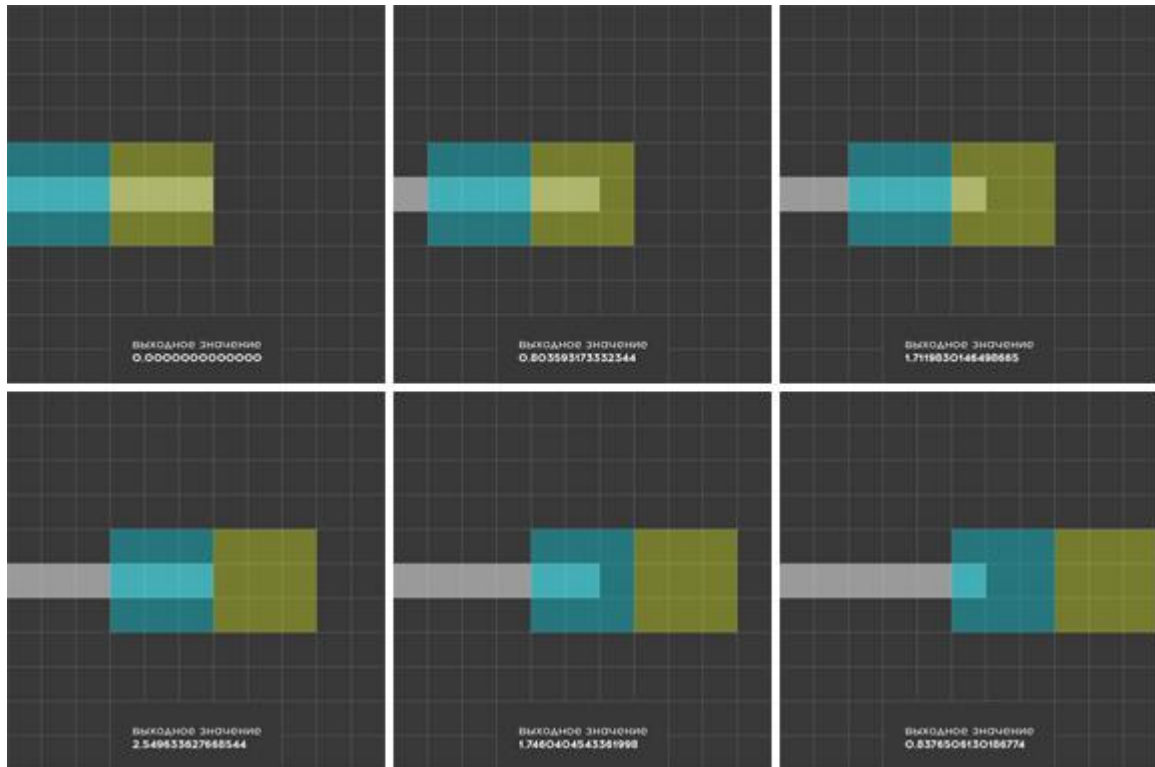


Рис. 7. Результат работы модели Хейтгера

Fig. 7. The result of the Heitger model

В неокогнитроне [19] для моделирования поведения нейронов конца линий использована схема с тормозным окружением в связях между S-клетками (простые нейроны) u_s и C-клетками (сложные нейроны) u_c .

$$u_c(x, y) = \psi[\sum_A a_c(n, m)u_s(x + n, y + m)], \quad (11)$$

где A – размер рецептивного поля,

$$\psi(x) = \frac{f(x)}{1+f(x)} \quad (12)$$

Матрица весов связей $a_c(n, m)$ формируется с помощью вейвлета «сомбреро». На рис. 8 приведено сравнение графиков вейвлета «мексиканская шляпа» и фильтра Габора, используемого нами.

$$a_c(n, m) = \exp\left(-\frac{n^2+m^2}{2\sigma^2}\right)\left(1 - \frac{n^2+m^2}{2\sigma^2}\right). \quad (13)$$

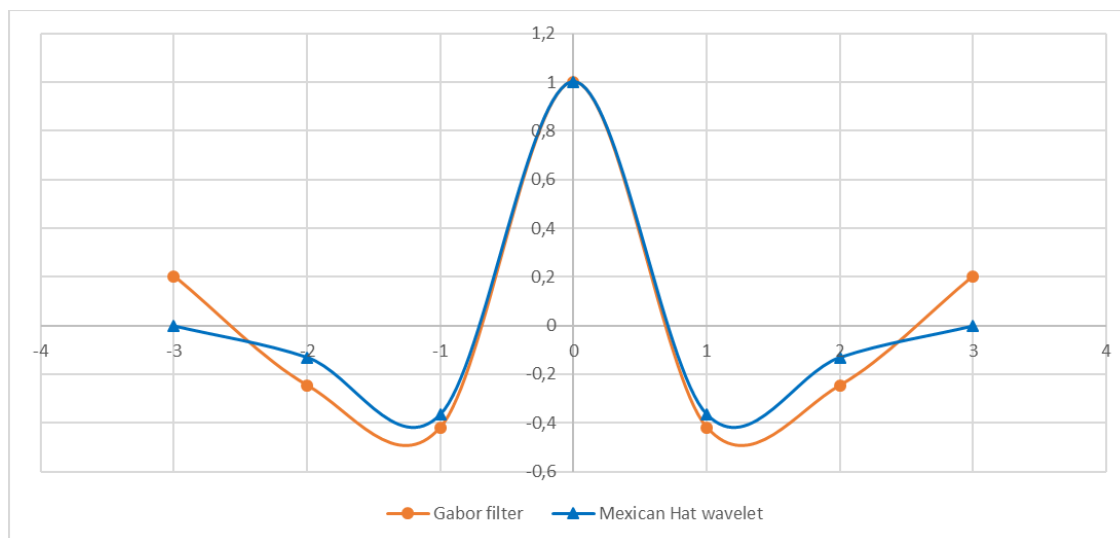


Рис. 8. Сравнение графиков фильтров
Fig. 8. Comparison of filter graphs

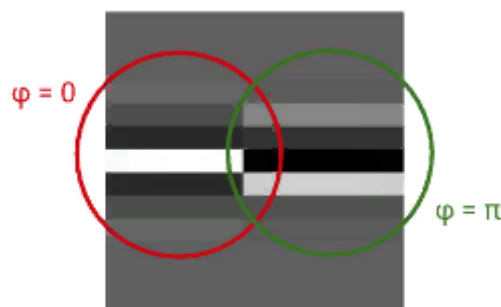
3. Нейрон конца линий

Мы предлагаем моделировать нейрон конца линий как сложную клетку (см. (14)), связанную с двумя близко расположенными простыми клетками, основанными на фильтре Габора, с одинаковой ориентацией θ , но находящимся в противофазе (рис. 9).

$$C_e = G_t(x', y', \varphi_1) + G_t(x' + \Delta x, y' + \Delta y, \varphi_2 = \varphi_1 + \pi), \quad (14)$$

где t – индекс типа простого нейрона ($=1,2$).

Рис. 9. Схема работы нейрона по модели CE
Fig. 9. The scheme of the neuron according to the CE model



Продолжение линии в пределах рецептивного поля нейрона, находящегося в противофазе к другому нейрону, приведет к уменьшению активации (рис. 10). Таким образом, используются данные от простых клеток выделения линий. Размер рецептивного поля нейрона конца линий также составляет 13×13 пикселей.

0.97	1.96	2.95	3.95	4.95	5.93	6.90	5.92	4.92	3.92	2.92	1.94	0.97

Рис. 10. Активация нейрона конца линий в зависимости от длины линии
 Fig. 10. Activation of the neuron of the end of the lines depending on the length of the line

Благодаря использованию фильтра Габора сохраняется чувствительность модели нейрона конца линий к изменению контраста.

4. Результаты

На рис. 11 представлены результаты работы *СЕ*-модели нейрона конца линий, предлагаемой нами. Так как в каждом пикселе изображения ищутся концы линий всех ориентаций, то на изображении показаны ориентации и выходные значения нейронов, получивших максимум активации.

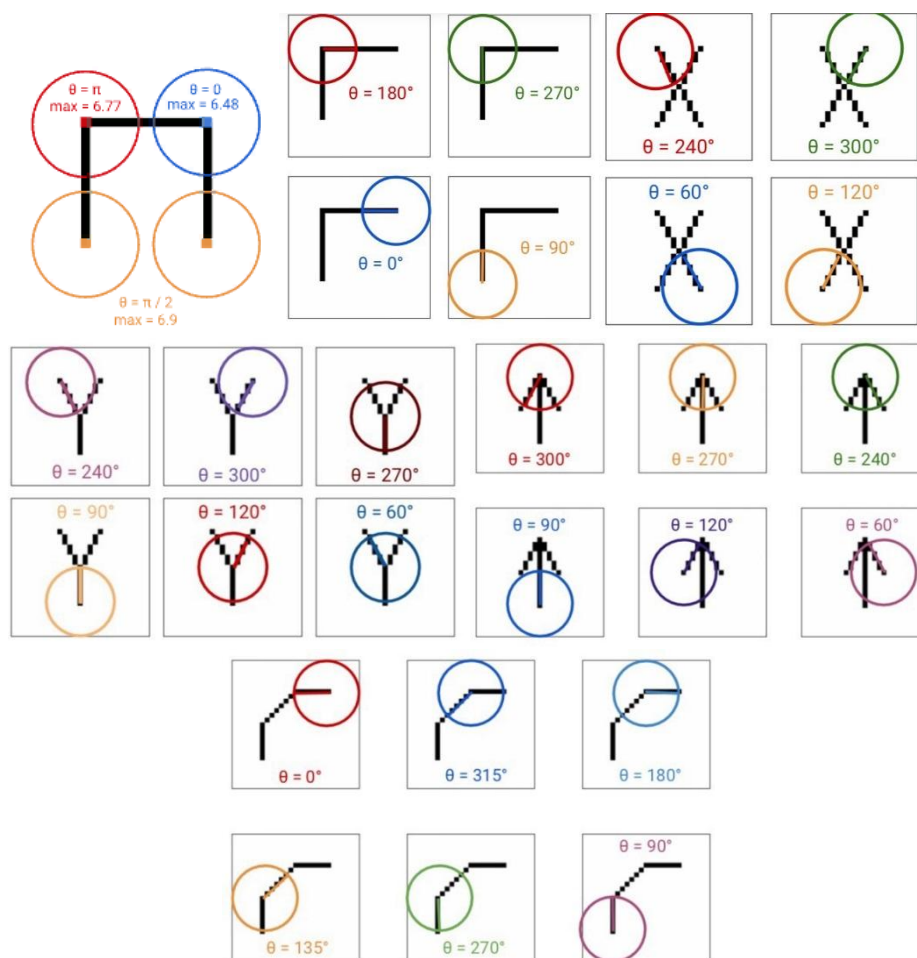


Рис. 11. Результаты работы модели нейрона
 Fig. 11. Results of the neuron model

Заключение

Как видно из рис. 6, $B-V$ модель не позволяет с точностью до пикселя выделять конец линии. Реакция нейрона конца линий в неокогнитроне отличается во втором знаке после запятой, что может нивелироваться колебаниями яркости пикселей на естественном изображении. Модель Хейтгера и предложенная нами модель CE близки и показывают высокую селективную способность, но модель Хейтгера более сложна в вычислительном плане. К тому же вычисление квадратурной пары фильтров Габора, необходимое для энергии Габора, оправданно для цифровой обработки сигналов, где выравнивает влияние низких частот. При работе с изображением в пространстве $CIE\ L^*a^*b^*$ в этом нет необходимости. Еще одним достоинством нашей модели является легкая интеграция со свёрточными нейронными сетями, где функция рецепторов базируется на использовании фильтра Габора.

Список литературы / References

1. Serre T., Wolf L., Bileschi S., Riesenhuber M., Poggio T. Robust Object Recognition with Cortex-Like Mechanisms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, vol. 29, no. 3, pp. 411–426. DOI 10.1109/TPAMI.2007.56
2. Díaz-Pernas F. J., Martínez-Zarzuela M., Antón-Rodríguez M., González-Ortega D. Learning and surface boundary feedbacks for colour natural scene perception. *Applied Soft Computing*, 2017, vol. 61, pp. 30–41. DOI 10.1016/j.asoc.2017.07.055
3. Hubel D. H. Eye, brain, and vision. New York, Scientific American Library: Distributed by W. H. Freeman, 1988.
4. Yazdanbakhsh A., Livingstone M. S. End stopping in V1 is sensitive to contrast. *Nature Neuroscience*, 2006, vol. 9, no. 5, pp. 697–702. DOI 10.1038/nn1693
5. Eskikand P. Z., Kameneva T., Ibbotson M. R., Burkitt A. N., Grayden D. B. A Possible Role for End-Stopped V1 Neurons in the Perception of Motion: A Computational Model. *PLOS ONE*, 2016, vol. 11, no. 10, p. e0164813. DOI 10.1371/journal.pone.0164813
6. Nicholls J. G. (ed.) From neuron to brain. 5th ed. Sunderland, Mass, Sinauer Associates, 2012.
7. Kugaevskikh A. V., Sogreshilin A. A. Analyzing the Efficiency of Segment Boundary Detection Using Neural Networks. *Optoelectron. Instrument. Proc.*, 2019, vol. 55, no. 4, pp. 414–422. DOI 10.3103/S8756699019040137
8. Gabor D. Theory of communication. *Journal of the Institution of Electrical Engineers*, 1946, vol. 93, no. 26, pp. 429–457.
9. Daugman J. G. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *J. Opt. Soc. Am. A*, 1985, vol. 2, no. 7, p. 1160. DOI 10.1364/JOSAA.2.001160
10. Petkov N. Biologically motivated computationally intensive approaches to image pattern recognition. *Future Generation Computer Systems*, 1995, vol. 11, no. 4–5, pp. 451–465. DOI 10.1016/0167-739X(95)00015-K
11. Kruizinga P., Petkov N. Nonlinear operator for oriented texture. *IEEE Transactions on Image Processing*, 1999, vol. 8, no. 10, pp. 1395–1407. DOI 10.1109/83.791965
12. Bovik A. C., Clark M., Geisler W. S. Multichannel texture analysis using localized spatial filters. *IEEE Trans. Pattern Anal. Machine Intell.*, 1990, vol. 12, no. 1, pp. 55–73. DOI 10.1109/34.41384
13. Pollen D. A., Ronner S. F. Visual cortical neurons as localized spatial frequency filters. *IEEE Trans. Syst., Man, Cybern.*, 1983, vol. SMC-13, no. 5, pp. 907–916. DOI 10.1109/TSMC.1983.6313086
14. Akbarinia A., Parraga C. A. Feedback and Surround Modulated Boundary Detection. *International Journal of Computer Vision*, 2018, vol. 126, no. 12, pp. 1367–1380. DOI 10.1007/s11263-017-1035-5

15. **Bhattacharjee S. K., Vandergheynst P.** End-stopped wavelets for detecting low-level features. In: SPIE's International Symposium on Optical Science, Engineering, and Instrumentation. Denver, CO, Oct. 1999, pp. 732–741. DOI 10.1117/12.366829
16. **Rodrigues J., J. Buf M. H. du.** Multi-scale keypoints in V1 and beyond: Object segregation, scale selection, saliency maps and face detection. *Biosystems*, 2006, vol. 86, no. 1–3, pp. 75–90. DOI 10.1016/j.biosystems.2006.02.019
17. **Würtz R. P., Lourens T.** Corner detection in color images through a multiscale combination of end-stopped cortical cells. *Image and Vision Computing*, 2000, vol. 18, no. 6–7, pp. 531–541. DOI 10.1016/S0262-8856(99)00061-X
18. **Heitger F., Rosenthaler L., Heydt R. von der, Peterhans E., Kübler O.** Simulation of neural contour mechanisms: from simple to end-stopped cells. *Vision Research*, 1992, vol. 32, no. 5, pp. 963–981. DOI 10.1016/0042-6989(92)90039-L
19. **Fukushima K.** Neocognitron for handwritten digit recognition. *Neurocomputing*, 2003, vol. 51, pp. 161–180. DOI 10.1016/S0925-2312(02)00614-8

Информация об авторе

Александр Владимирович Кугаевских, кандидат технических наук

Information about the Author

Alexander V. Kugaevskikh, Candidate of Sciences (Techniques)

*Статья поступила в редакцию 10.06.2021;
одобрена после рецензирования 01.08.2021; принята к публикации 01.08.2021
The article was submitted 10.06.2021;
approved after reviewing 01.08.2021; accepted for publication 01.08.2021*

Научная статья

УДК 004.021

DOI 10.25205/1818-7900-2021-19-3-61-69

Методы Paper2vec и Cite2vec для анализа коллекций научных публикаций

Николай Игоревич Тихонов

Новосибирский государственный университет
Новосибирск, Россия
t.kolya54@gmail.com, <https://orcid.org/0000-0001-8765-3263>

Аннотация

Визуализации коллекций научных публикаций используются для лучшего понимания наборов данных. При построении таких визуализаций могут использоваться различные методы анализа текстовых коллекций. В данной статье речь идет о двух методах анализа – Paper2vec и Cite2vec, которые в своей работе используют информацию о цитировании и получают векторные представления документов. Чтобы продемонстрировать работу методов, были разработаны визуализации, которые описаны в данной работе.

Ключевые слова

визуализация коллекций документов, векторное представление документов, сети цитирования, контексты цитирования

Благодарности

Работа выполнена под научным руководством З. В. Апанович

Для цитирования

Тихонов Н. И. Методы Paper2vec и Cite2vec для анализа коллекций научных публикаций // Вестник НГУ. Серия: Информационные технологии. 2021. Т. 19, № 3. С. 61–69. DOI 10.25205/1818-7900-2021-19-3-61-69

Paper2vec and Cite2vec Methods for Analyzing Collections of Scientific Publications

Nikolay I. Tikhonov

Novosibirsk State University
Novosibirsk, Russian Federation
t.kolya54@gmail.com, <https://orcid.org/0000-0001-8765-3263>

Abstract

Visualizations are used to better understand collections of scientific publications. Various methods of analyzing text collections can be used to build these visualizations. This article discusses two methods Paper2vec and Cite2vec that get vector representations of documents using citation information. To demonstrate a work of these techniques and an example of their application, visualizations were developed, which are described in this paper.

Keywords

visualization of document collections, vector representation of documents, citation networks, citation contexts

Acknowledgements

The work performed under the guidance of the scientific supervisor Z. V. Apanovich

For citation

Tikhonov N. I. Paper2vec and Cite2vec Methods for Analyzing Collections of Scientific Publications. *Vestnik NSU. Series: Information Technologies*, 2021, vol. 19, no. 3, p. 61–69. (in Russ.) DOI 10.25205/1818-7900-2021-19-3-61-69

© Тихонов Н. И., 2021

Введение

Коллекции научных публикаций растут быстрыми темпами. Ученым доступны порталы, содержащие большое количество документов. Например, сайт SpringerLink¹ в данный момент содержит 14,07 млн ресурсов. Исследование такого большого объема данных – трудоемкий процесс. Для сокращения трудозатрат, поиска нужных / схожих документов, оценки научного вклада определенных публикаций и обнаружения скрытых связей между документами применяются методы визуализации документов.

В основе методов визуализации документов могут лежать различные модели представления документов [1]. В последние годы чрезвычайно популярны методы векторных представлений слов для обработки естественных языков. Вслед за ними стали появляться методы анализа текстовых коллекций для получения векторных представлений документов.

Хотя существует множество систем анализа документов, новые методы могут вносить новое понимание коллекций, иметь большую производительность при анализе больших коллекций документов, находить новые связи между документами.

В данной статье идет речь о двух методах анализа – Paper2vec и Cite2vec, которые в своей работе используют информацию о цитировании и получают векторные представления документов. Приведены пример использования данных представлений для построения визуализаций и описание двух рассматриваемых методов. Также описаны наши эксперименты с этими двумя методами, включающие визуализацию результатов работы методов и описание проблем при визуализации такого вида информации.

1. Методы анализа текстовых коллекций

Вслед за методами векторных представлений слов для обработки естественного языка [2–5] стали появляться методы для векторного представления текстовых документов.

Рассматриваемые два метода используют информацию о цитировании: первый – контексты цитирования, второй – информацию о том, кто кого цитирует в коллекции.

1.1. Paper2vec

Paper2vec [6] использует граф цитирования для получения векторных представлений документов, не требуя полный текст документов или контексты цитирования. Это позволяет применять его в базах данных, где полный текст или контексты цитирования не поддерживаются.

Алгоритм Paper2vec

Требует: база данных научных работ D , содержащих отношения цитирования;

Выдает: векторные представления W исследовательских работ, содержащихся в D ;

- 1) построить сеть цитирования из D ;
- 2) построить матрицу весов отношений цитирования исследовательских работ из построенной сети;
- 3) стохастически минимизировать функцию стоимости для получения векторов документов W ;
- 4) вернуть W .

Набор соседних узлов в сети цитирования рассматривается авторами как «контекст цитирования», при этом считается, что контекст не ограничивается отношениями прямого цитирования, косвенные цитаты тоже несут в себе семантику рассматриваемого документа, но

¹ <https://link.springer.com/>

уже с меньшим весом. Таким образом, на втором шаге алгоритма определяются контексты цитирования, и задается весовая схема на сети цитирования, которая учитывает следующие характеристики:

- цитирующие и цитируемые статьи вместе помогают предсказать содержимое текущего документа;
- статьи, которые ссылаются косвенно, не так актуальны, как статьи, которые ссылаются непосредственно (т. е. имеют меньший вес);
- транзитивность. Предположим, что статья А цитирует статью В, а статья В цитирует статью С: чем меньше вес отношения цитирования между А и В или В и С, тем меньше вес отношения цитирования между А и С.

Чтобы удовлетворить вышеуказанным свойствам, весовая схема основана на вероятности случайного блуждания. Данные веса участвуют в функции стоимости, которая стохастически минимизируется.

Эксперименты по оценке работы метода проводились на основе CITREC, открытой системы оценки мер сходства на основе цитирования, которая предоставляет наборы научных данных, базовые измерения и некоторые реализации предыдущих алгоритмов, основанных на цитировании. Набор данных CITREC включает в себя коллекции PubMed Central Open Access Subset (PMCOS) и TREC Genomics. Отношения цитирования были извлечены из текста документов при помощи методов, предоставленных CITREC.

Paper2vec сравнивался с Amsler, CPA, DeepWalk (рис. 1). Каждый метод для каждого документа определял K похожих документов. Эти данные сравнивались с эталонными, и определялся средний коэффициент пересечения. По результатам экспериментов, Paper2vec и DeepWalk значительно превосходят модели Amsler и CPA. Paper2vec лучше, чем DeepWalk на малых значениях K , что означает, что он может найти лучшие результаты в первых нескольких документах.

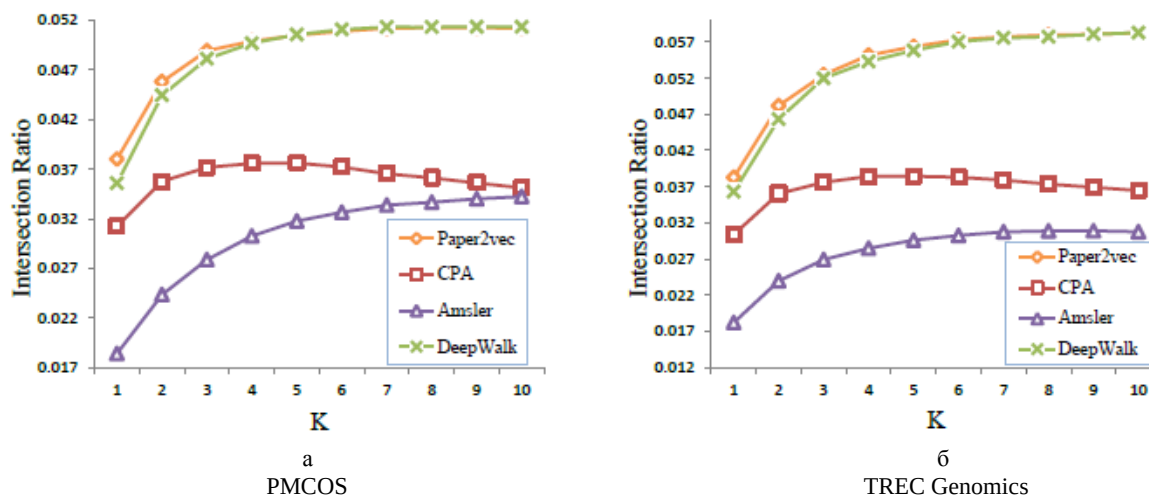


Рис. 1. Сравнение метода Paper2vec на двух наборах данных (PMCOS и TREC Genomics) с Amsler, CPA и DeepWalk

Fig. 1. Comparison of the Paper2vec method on two datasets (PMCOS and TREC Genomics) with Amsler, CPA and DeepWalk

1.2. Cite2vec

Система Cite2vec [7] предназначена для векторного представления документов и их динамического исследования. Для обучения векторных моделей используются контексты цитирования научных публикаций.

Для представления документов в векторном пространстве Cite2vec использует дополненную модель Skip-gram из Word2vec. Модель Skip-gram принимает последовательность слов и назначает точку в многомерном пространстве каждому слову, так чтобы слова, которые встречаются вместе в предложениях, были близки в пространстве. Разработчики Cite2vec дополняют эту модель словарем документов, которые цитируются в рассматриваемой коллекции, чтобы исследовать слова и документы в одном и том же пространстве.

Каждый документ, упоминаемый в списках цитируемой литературы, получает в Cite2vec уникальный идентификатор. Таким образом, с каждым документом сопоставляется уникальное слово, что позволяет представлять и слова, и документы в едином векторном пространстве. При этом векторы, соответствующие словам-идентификаторам статей, оказываются в окружении векторов-слов, используемых авторами цитирующих публикаций для описания этих документов.

Полученные таким образом векторы слов и документов обладают рядом полезных характеристик. Во-первых, слова несут семантику тем документов. Кроме того, единое пространство наследует линейную структуру, это позволяет связывать слова и документы при помощи простой арифметики.

Система включает в себя визуализацию и позволяет пользователю динамически просматривать документы на основе информации о том, как их используют остальные документы. Пользователь имеет возможность сформулировать любое понятие, состоящее из нескольких слов. При выборе такого понятия все слова, входящие в понятие, суммируются со словами, присутствующими в визуализации, и исследуемый документ перемещается ближе к тому слову, чья сумма со словами выбранного понятия наиболее точно отражает смысл рассматриваемого документа. В качестве понятий можно указывать сами документы, они несут в себе богатый набор значений.

2. Эксперименты

Рассмотрим пример визуализации полученных результатов анализа текстов и визуально оценим работу двух указанных методов на подготовленном наборе данных. Опишем проблемы, возникающие при визуализации векторов большой размерности.

2.1. Набор данных

Данные для экспериментов были взяты с сайта CiteSeerX² и содержат 3 811 контекстов цитирования из 363 статей. Статьи были отобраны поиском по месту публикации, содержащим слово *visualization*. Информация о каждой статье содержит: уникальный идентификатор статьи DOI, уникальные идентификаторы цитируемых статей и контексты цитирования, название, автора, ключевые слова и место публикации.

Далее были подготовлены входные данные для каждого из методов (Paper2vec, Cite2vec) и получены векторные представления документов, а также, в случае метода Cite2vec, векторные представления слов из контекстов цитирования.

Далее подготовка данных визуализации для каждого метода имеет различия. На полученных векторах документов метода Paper2vec применялись сначала иерархическая кластериза-

² <https://citeseerx.ist.psu.edu/>

ция (чтобы сохранить часть семантики полученных векторных представлений), а затем T-SNE [8] для уменьшения размерности до двух.

В случае метода Cite2vec сначала находились репрезентативные слова документов, которые отражают семантику тем, аналогично подходу оригинальной статьи [7]. А именно в цикле на каждой итерации определялся самый дальний документ от уже обработанных и для него находилось ближайшее слово с некоторыми ограничениями. Если ближайшее уже выбрано или расстояние до уже выбранных меньше некоторого порогового значения, то отбрасываем это слово и рассматриваем следующее ближайшее. При этом число слов ограничено сверху, чтобы выбранные слова были значимы для пользователя. Затем для выбранных слов применялся метод T-SNE, и все документы рассеивались в двумерном пространстве около самого близкого репрезентативного слова. Все расстояния считались по косинусной мере.

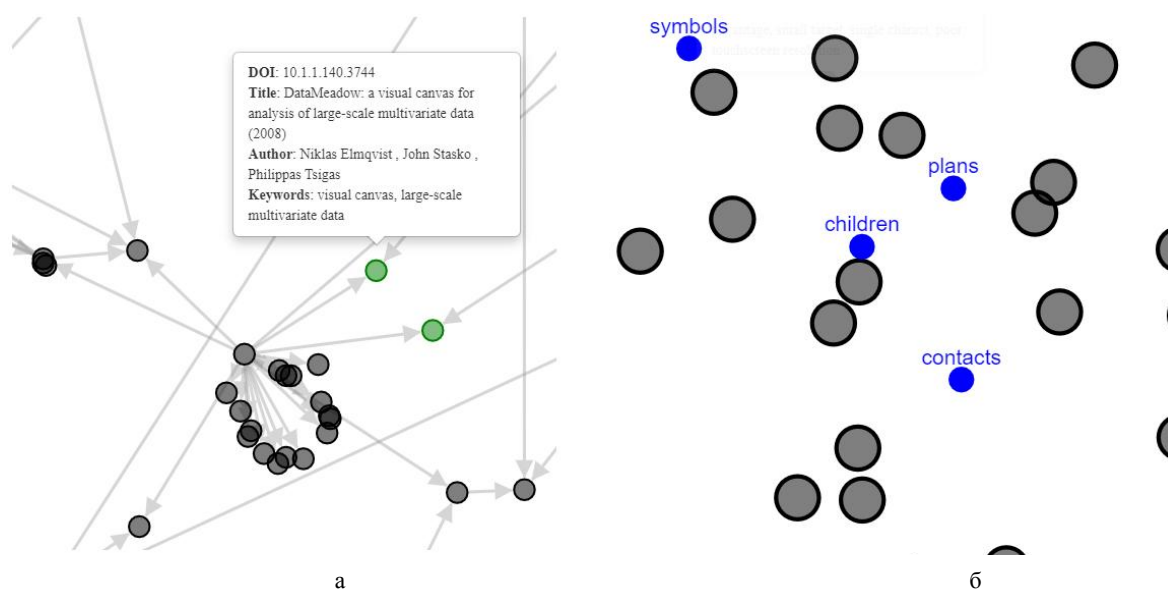


Рис. 2. Визуализация документов:

a – на примере Paper2vec (в виде прозрачных дисков, описание документа во всплывающем окне, ребрами представлены отношения цитирования); *б* – на примере Cite2vec (ребра скрыты)

Fig. 2. Visualization of documents:

a – using the visualization Paper2vec as an example (in the form of transparent disks, document description in a pop-up window, the edges represent citation relationships);
b – using the visualization Cite2vec as an example (the edges are hidden)

2.2. Визуализация

Результаты обоих методов (Paper2vec и Cite2vec) визуализировались при помощи JavaScript библиотеки Cytoscape³. Визуализации обоих методов имеют как сходства, так и различия в функционале.

На обоих изображениях каждый документ представляется в виде прозрачного диска (рис. 2, *a*), чтобы хорошо были видны наложения одного документа на другой. В визуализации Cite2vec слова показываются в виде непрозрачной точки синего цвета меньшего размера, возле точки отображается само слово (рис. 2, *б*)

Диски документов Paper2vec из одного и того же кластера окрашиваются в одинаковый цвет. Следует отметить, что из-за аппроксимационной природы t-SNE документы из одного кластера не обязательно размещаются близко друг к другу на 2D-карте.

³ <https://js.cytoscape.org/>

Пользователю доступен базовый набор методов взаимодействия с визуализациями, который позволяет перемещать точки, масштабировать изображение, смотреть описания документов во всплывающих окнах, выполнять поиск, и управлять выводом дополнительных элементов визуализации. Кроме этого в Paper2vec доступна настройка кластеризации.

2.2.1 Поиск

Чтобы быстрее находить интересные документы, был добавлен поиск статей по DOI, автору, названию, ключевым словам, месту публикации (рис. 3, а) и, в случае Cite2vec, по репрезентативным словам документов. Можно настраивать, где именно производится поиск. В результатах поиска показывается информация (DOI, название, автор, место публикации, ключевые слова, ссылка на статью в CiteSeerX) из подходящих документов с возможностью выделить все найденные документы (документы и слова в визуализации Cite2vec) на карте либо один, выбранный из списка результатов.

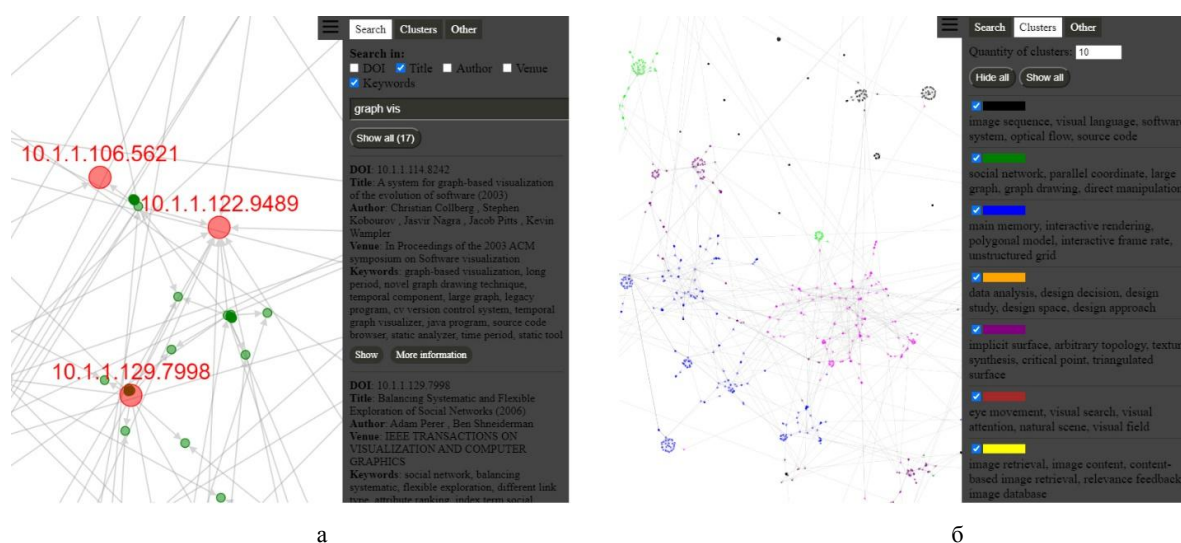


Рис. 3. (а) Поиск статей в визуализации. Выделены статьи из списка результата поиска «graph vis» в названиях документов и ключевых словах. На примере визуализации Paper2vec.

(б) Иерархическая кластеризация векторов Paper2vec. Количество кластеров выбрано равное десяти
Fig. 3. (a) Search articles in visualization. Selected articles from the list of “graph vis” search results in document titles and keywords. Using the visualization Paper2vec as an example.

(b) Hierarchical clustering of Paper2vec vectors. The number of clusters was chosen equal to ten.

2.2.2 Кластеры

Для сохранения исходной структуры векторов Paper2vec применяется иерархическая кластеризация до применения метода снижения размерности T-SNE. На изображении каждый кластер окрашен в уникальный цвет, можно скрывать / показывать каждый из кластеров, управлять их количеством (рис. 3, б). В настройках кластеризации возле каждого кластера отображаются пять самых часто встречаемых ключевых слов среди его документов, исключая те, которые содержатся более чем в $\frac{1}{3}$ кластеров.

2.2.3 Дополнительные элементы визуализации

Для явного отражения семантики данных используются дополнительные элементы визуализации.

Отношения цитирования в обеих визуализациях показаны ориентированными ребрами от цитирующей статьи к цитируемой (см. рис. 2, *a*). Вы можете в любой момент отключить показ ребер, если они загромождают изображение. Кроме этого в визуализации Cite2vec дополнительными элементами являются слова, которые также можно скрыть.

2.3 Проблемы при визуализации векторов большой размерности

Векторные представления коллекций научных публикаций могут быть очень эффективными, но их трудно интерпретировать. Методы уменьшения размерности теряют множественные признаки схожести / разности векторов, поэтому нужны дополнительные средства визуализации, которые позволяли бы лучше ухватывать семантику этих представлений, и, значит, улучшать эти представления.

В визуализации Cite2vec используются репрезентативные слова документов, которые несут семантику тем документов. В визуализации Paper2vec применяется иерархическая кластеризация перед уменьшением размерности, чтобы сохранить часть семантики векторных представлений.

2.4 Вывод

На рис. 4 представлены визуализации результатов методов Paper2vec и Cite2vec. В Cite2vec слова и документы моделируются в едином векторном пространстве, что позволяет добавить дополнительные элементы визуализации (документы в окружении слов, которые несут семантику тем документов; возможность связывать слова и документы при помощи простой арифметики). Для лучших результатов нужны большие наборы данных. В статье [7] авторами был показан функционал для ввода ключевых фраз. После их ввода документы перемещаются к их наиболее подходящей композиции репрезентативных слов и введенных пользователем фраз.

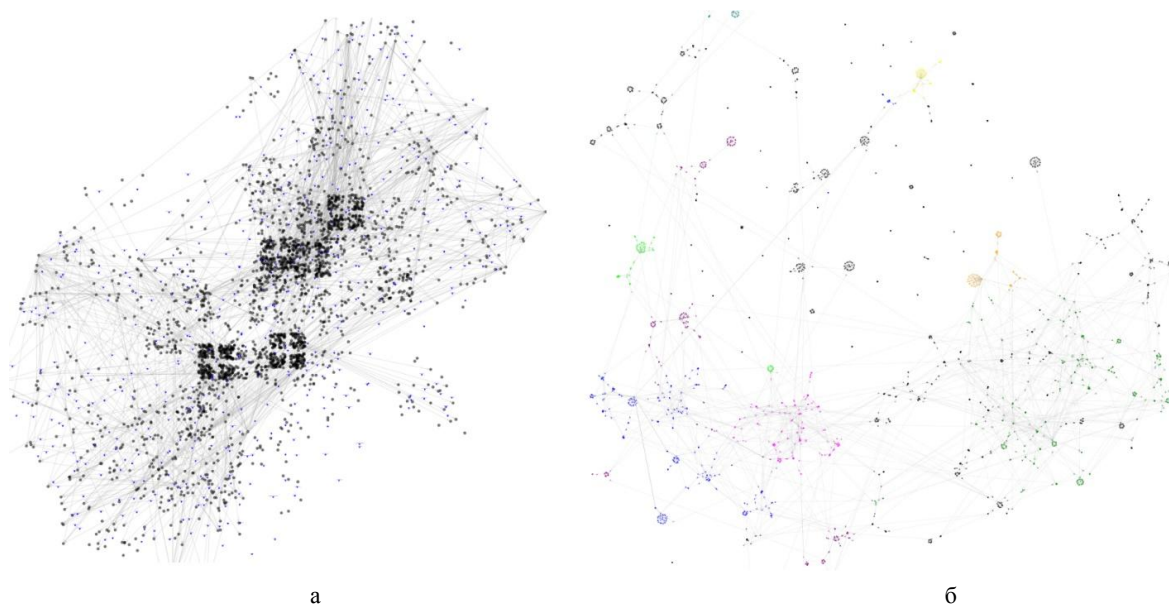


Рис. 4. Визуализация результатов: *a* – Cite2vec; *б* – Paper2vec
Fig. 4. Visualization of results: *a* – Cite2vec; *b* – Paper2vec

Paper2vec получает представления для цитирующих и цитированных документов, в отличие от Cite2vec, где только документы из контекстов цитирования (только цитированные). В Paper2vec цитирующие и цитируемые статьи, как правило, расположены близко. На изображении отчетливо видно множество небольших скоплений документов.

Заключение

В данной работе представлены два современных метода анализа текстовых коллекций – Paper2vec и Cite2vec, которые получают представления документов (и слов из контекстов цитирования в случае Cite2vec) в векторном пространстве. Для Paper2vec не важен язык, на котором написаны статьи, и он применим к базам данных, которые не поддерживают полные тексты документов, метод получает векторные представления для цитируемых и цитирующих документов. Для лучшей работы Cite2vec нужны большие наборы данных, данный метод получает векторные представления слов и документов в едином векторном пространстве, при этом документы оказываются в окружении слов, которые авторы используют в контекстах цитирования для описания данных документов.

Чтобы продемонстрировать работу методов были созданы визуализации результатов работы каждого метода. При визуализации такого вида информации (многомерных векторов) возникают проблемы с потерей семантики данных представлений в связи с уменьшением размерности. В статье были описаны примеры того, как можно сохранить часть семантики.

Список литературы

1. **Апанович З. В.** Эволюция методов визуализации коллекций научных публикаций // Электронные библиотеки. 2018. Т. 21, № 1. С. 2–42.
2. **Mikolov T., Sutskever I., Chen K., Corrado G. S., Dean J.** Distributed Representations of Words and Phrases and Their Compositionality. *Advances in Neural Information Processing Systems*, 2013, vol. 26, pp. 3111–3119.
3. **Pennington J., Socher R. D., Manning C.** Glove: Global vectors for word representation. In: *Proceedings of the Empirical Methods in Natural Language Processing (EMNLP 2014)*, 2014, pp. 1532–1543. DOI 10.3115/v1/D14-1162
4. **Bojanowski P., Grave E., Joulin A., Mikolov T.** Enriching Word Vectors with Subword Information. *Transactions of the Association for Computational Linguistics*, 2017, vol. 5, pp. 135–146. DOI 10.1162/tac1_a_00051
5. **Peters M., Neumann M., Iyyer M., Gardner M., Clark C., Lee K., Zettlemoyer L.** Deep contextualized word representations. In: *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2018, vol. 1, pp. 2227–2237. DOI 10.18653/v1/N18-1202
6. **Tian H., Zhuo H. H.** Paper2vec: Citation-Context Based Document Distributed Representation for Scholar Recommendation. *ArXiv*. abs/1703.06587, 2017.
7. **Berger M., McDonough K., Seversky Lee M.** Cite2vec: Citation-Driven Document Exploration via Word Embeddings. *IEEE Transactions on Visualization and Computer Graphics*, 2017, vol. 23, no. 1, pp. 691–700. DOI 10.1109/TVCG.2016.2598667
8. **Maaten L. van der, Hinton G.** Visualizing data using t-SNE. *Journal of Machine Learning Research*, 2008, vol. 9, pp. 2579–2605.

References

1. **Apanovich Z. V.** Evolution of Visualization Methods for Research Publication Collections. *Elektronnye biblioteki*, 2018, vol. 21, no. 1, pp. 2–42. (in Russ.)

2. **Mikolov T., Sutskever I., Chen K., Corrado G. S., Dean J.** Distributed Representations of Words and Phrases and Their Compositionality. *Advances in Neural Information Processing Systems*, 2013, vol. 26, pp. 3111–3119.
3. **Pennington J., Socher R. D., Manning C.** Glove: Global vectors for word representation. In: *Proceedings of the Empirical Methods in Natural Language Processing (EMNLP 2014)*, 2014, pp. 1532–1543. DOI 10.3115/v1/D14-1162
4. **Bojanowski P., Grave E., Joulin A., Mikolov T.** Enriching Word Vectors with Subword Information. *Transactions of the Association for Computational Linguistics*, 2017, vol. 5, pp. 135–146. DOI 10.1162/tacl_a_00051
5. **Peters M., Neumann M., Iyyer M., Gardner M., Clark C., Lee K., Zettlemoyer L.** Deep contextualized word representations. In: *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2018, vol. 1, pp. 2227–2237. DOI 10.18653/v1/N18-1202
6. **Tian H., Zhuo H. H.** Paper2vec: Citation-Context Based Document Distributed Representation for Scholar Recommendation. *ArXiv. abs/1703.06587*, 2017.
7. **Berger M., McDonough K., Seversky Lee M.** Cite2vec: Citation-Driven Document Exploration via Word Embeddings. *IEEE Transactions on Visualization and Computer Graphics*, 2017, vol. 23, no. 1, pp. 691–700. DOI 10.1109/TVCG.2016.2598667
8. **Maaten L. van der, Hinton G.** Visualizing data using t-SNE. *Journal of Machine Learning Research*, 2008, vol. 9, pp. 2579–2605.

Информация об авторе

Николай Игоревич Тихонов, аспирант

Information about the Author

Nikolay I. Tikhonov, Graduate Student

Статья поступила в редакцию 20.07.2021;

одобрена после рецензирования 21.08.2021; принята к публикации 21.08.2021

The article was submitted 20.07.2021;

approved after reviewing 21.08.2021; accepted for publication 21.08.2021

Научная статья

УДК 004.9, 624.131

DOI 10.25205/1818-7900-2021-19-3-70-82

Разработка методики перехода к технологиям информационного моделирования в инженерных изысканиях (на примере инженерно-экологических изысканий)

**Антон Сергеевич Флеенко¹
Александр Федорович Демьяненко²**

^{1,2} Российский университет транспорта (РУТ – МИИТ)
Москва, Россия

¹ fleenkospb@mail.ru, <https://orcid.org/0000-0002-2829-9361>

² afdco89@rambler.ru, <https://orcid.org/0000-0002-2859-7852>

Аннотация

Технология информационного моделирования становятся одним из эффективных инструментов работы с данными. Однако до сих пор не существует конкретных рекомендаций по переходу к использованию цифровых моделей на этапе изысканий. Целью работы выступает поиск и оценка путей интеграции процедур информационного моделирования с процессами, протекающими в системах инженерных изысканий и проектирования. В статье разработана и охарактеризована с применением методов системного анализа, форсайт-исследования и риск-менеджмента методика модернизации инженерных изысканий в указанном контексте. Изменения связей, состава и результатов, описание этапов внедрения рассмотрены на примере инженерно-экологического блока. Авторы считают, что на фоне формирования нормативной правовой базы информационного моделирования в России наиболее результативным направлением является представление итога инженерных изысканий в виде стандартных технических отчетов, дополненных цифровыми моделями по видам изысканий.

Ключевые слова

информационное моделирование, цифровая информационная модель, инженерные изыскания, инженерно-экологические изыскания, системный анализ, риск-менеджмент, форсайт-исследование

Для цитирования

Флеенко А. С., Демьяненко А. Ф. Разработка методики перехода к технологиям информационного моделирования в инженерных изысканиях (на примере инженерно-экологических изысканий) // Вестник НГУ. Серия: Информационные технологии. 2021. Т. 19, № 3. С. 70–82. DOI 10.25205/1818-7900-2021-19-3-70-82

© Флеенко А. С., Демьяненко А. Ф., 2021

ISSN 1818-7900 (Print). ISSN 2410-0420 (Online)

Вестник НГУ. Серия: Информационные технологии. 2021. Том 19, № 3. С. 70–82

Vestnik NSU. Series: Information Technologies, 2021, vol. 19, no. 3, pp. 70–82

Development of a Methodology for the Transition to Information Modeling Technologies in Engineering Surveys (On the Example of Engineering and Environmental Surveys)

Anton S. Fleenko¹, Alexander F. Demianenko²

^{1,2} Russian University of Transport (MIIT)
Moscow, Russian Federation

¹ fleenkospb@mail.ru, <https://orcid.org/0000-0002-2829-9361>

² afdco89@rambler.ru, <https://orcid.org/0000-0002-2859-7852>

Abstract

Information modeling technology also known as BIM is becoming an important part of working with data in various fields of activity. At the same time, there are still no specific recommendations for the transition to the use of digital models at the survey stage. Although engineering survey data is the basis for further actions within the life cycle of objects. The article characterizes the methodology for the modernization of engineering surveys in the context of information modeling, operating with the methods of systems analysis, foresight research and risk management. The combination of these methods allows you to identify ways of development and characterize them. All changes in the engineering survey system are described using the example of the engineering and environmental section. This is due to the fact that environmental engineering surveys use information from other sections of the survey and show reactions to changes in the system. The authors believe that under existing conditions, the most effective direction is the presentation of the results of engineering surveys in a standard form, provided that they are supplemented with a set of digital models by type of survey. In this case, it is assumed that the role of information modeling will gradually increase.

Keywords

information modeling, digital information model, engineering surveys, engineering and environmental surveys, system analysis, risk management, foresight research

For citation

Fleenko A. S., Demianenko A. F. Development of a Methodology for the Transition to Information Modeling Technologies in Engineering Surveys (On the Example of Engineering and Environmental Surveys). *Vestnik NSU. Series: Information Technologies*, 2021, vol. 19, no. 3, p. 70–82. (in Russ.) DOI 10.25205/1818-7900-2021-19-3-70-82

Используемые термины

ВМ-проектирование (building information modeling, BIM, технология информационного моделирования) – совокупность инструментов, процессов и технологий, благодаря которым формируются компоненты цифровой модели, использующейся в процессе строительства с целью экономии финансов и времени, а также упрощения процессов за счет одновременной работы и визуализации объекта [1].

Цифровая информационная модель объекта – совокупность взаимосвязанных сведений, документов и материалов об объекте, формируемых в электронном виде на этапах выполнения инженерных изысканий, осуществления архитектурно-строительного проектирования, строительства, реконструкции, капитального ремонта, эксплуатации и (или) сноса объекта капитального строительства¹.

Инженерная цифровая модель местности – совокупность взаимосвязанных инженерно-геодезических, инженерно-геологических, инженерно-гидрометеорологических, инженерно-экологических, инженерно-геотехнических данных и данных о территории объекта, представленных в цифровом виде для автоматизированного решения задач управления процессами².

¹ СП 333.1325800.2020 «Информационное моделирование в строительстве. Правила формирования информационной модели объектов на различных стадиях жизненного цикла».

² Там же.

*Модель данных экологических наблюдений – цифровая модель, содержащая в своем составе данные инженерно-экологических изысканий*³.

Введение

Технологии информационного моделирования подразумевают представление данных в единой комплексной автоматизированной цифровой среде, часто визуализированной в трехмерном виде. В цифровой информационной модели с обширной базой данных, характеризующей объект и его составные части, возможна одновременная работа нескольких специалистов по ее изменению и дополнению. Информационное моделирование (далее ИМ), зародившись в строительной сфере, исходно было связано с системами, состоящими из виртуальных аналогов реальных элементов зданий. В настоящее время сфера применения данной технологии всё больше расширяется – цифровые модели начинают использоваться в качестве инструмента углубленной систематизации информации и повышения эффективности разработки и реализации любых проектов.

Авторы статьи полагают весьма перспективным применение ИМ в области инженерных изысканий в целом и инженерно-экологических изысканий в частности. Вместе с тем приходится констатировать, что ни отечественных, ни зарубежных исследований в этой области применения ИМ практически нет. В связи с этим заявляемая в этой статье цель научной работы, обозначаемая как поиск и оценка путей интеграции процедур ИМ с процессами, протекающими в системах инженерных изысканий и проектирования, представляется актуальной.

Разработка комплексной модели инженерных изысканий и ее развитие с учетом внедрения технологии ИМ

Для достижения поставленной цели использовались методы системного анализа: композиция и декомпозиция, моделирование и последующая оценка разработанных моделей. Процедура изысканий, в частности инженерно-экологических, сложна по структуре и функциям. Описание системы инженерных изысканий (далее системы) может быть осуществлено с помощью построения комплексной модели, включающей в себя характеристику внешней среды, входных и выходных связей (в рамках модели «черного ящика»), а также состава и структуры системы.

В качестве внешней среды по отношению к системе рассматриваются мегасистема организации, инициирующей процесс изысканий, и окружающая среда, представленная природными компонентами и социально-экономическими образованиями, такими как общественные организации, муниципальные образования и государственные структуры. Мегасистема организации представляет собой проектно-изыскательское предприятие, целевой функцией которого является разработка мегапроекта – проектных решений, содержащих в своем составе полный спектр информации: от изыскательской до технико-технологической и сметно-финансовой.

Из внешней среды в систему поступают средства производства, финансовые, трудовые и информационные ресурсы. Ограничениями для работы системы выступают природно-климатические особенности территории изысканий (сроки проведения изысканий), правовые (нормативно-правовые акты, регламентирующие изыскательскую деятельность) и технические (согласованное задание на проектирование, устанавливающее сроки и объемы работ) условия производства работ.

³ Требования к информационным моделям капитального строительства. Часть 6. Требования к представлению результатов инженерных изысканий для прохождения экспертизы при использовании технологии информационного моделирования // ГАУ г. Москвы «Московская государственная экспертиза». Редакция 1.0. М., 2019. 54 с.

Результатом работы системы является проектно-сметная информация по инженерному обеспечению мегапроекта. Это итоговая документация в виде технических отчетов и графических материалов по видам инженерных изысканий. Таким образом, основной выход системы как часть мегапроекта оказывается ее входом в мегасистему.

Определение состава и структуры инженерных изысканий осуществляется на основе теории организаций [2] и нормативно-правовых документов в области инженерных изысканий⁴. Система состоит из двух крупных модулей: функционального (предназначен для выполнения целевой функции системы – разработки проекта) и обеспечивающего (подразумевает гарантию ее выполнения посредством предоставления необходимых материально-технических, финансовых, инструментальных ресурсов и принятия организационно-управленческих решений). В рамках данного исследования наибольший интерес представляет функциональный модуль системы (табл. 1, рис. 1).

Таблица 1

Состав и характеристика функционального модуля системы инженерных изысканий

Table 1

Composition and characteristics of the functional module of the engineering survey system

Подсистема	Функция	Производные
Инженерно-геодезические изыскания (далее ИГДИ)	Получение топографо-геодезических данных и информации о ситуации и рельефе местности	Рельеф и геоположение
Инженерно-геологические и инженерно-геотехнические изыскания (далее ИГИ / ИГТИ)	Получение информации об инженерно-геологических и инженерно-геотехнических особенностях территории	Геологические, гидрогеологические и геотехнические данные
Инженерно-гидрометеорологические изыскания (далее ИГМИ)	Получение информации о гидрометеорологических особенностях территории	Гидрологические, метеорологические данные
Инженерно-экологические изыскания (далее ИЭИ)	Получение информации о состоянии окружающей среды и возможных источниках загрязнения	Данные экологических наблюдений

Описание абиотической составляющей – рельефа, геологических, геотехнических и гидрологических особенностей территории изысканий – отнесено к функциям подсистем ИГДИ, ИГИ / ИГТИ и ИГМИ. Сведения о характере рельефа и пространственном положении объекта исследования – производные ИГДИ – используются при разработке технической документации другими подсистемами функционального модуля. Отличительной чертой ИЭИ является применение в процессе изучения абиотической и биотической составляющей производных всех указанных подсистем.

⁴ СП 47.13330.2016 «Инженерные изыскания для строительства. Основные положения».

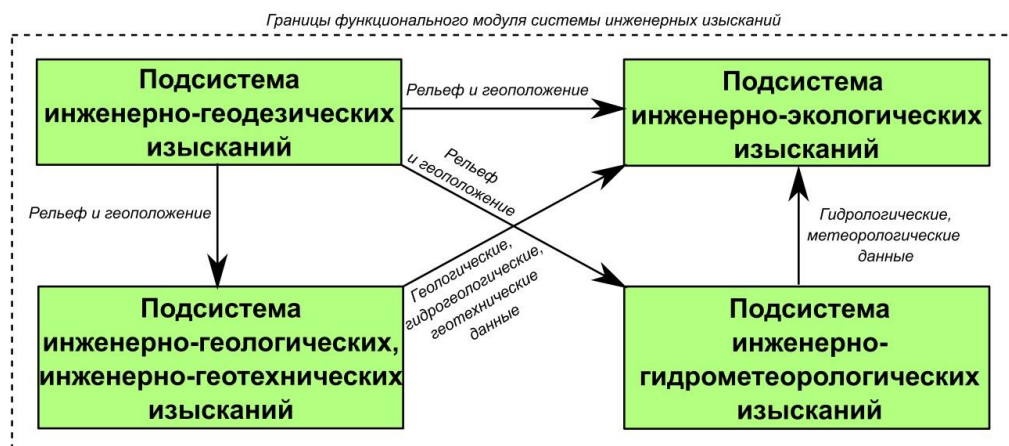


Рис. 1. Структура функционального модуля системы инженерных изысканий

Fig. 1. The structure of the functional module of the engineering survey system

На основе изложенного выше разработана комплексная модель инженерных изысканий (рис. 2). Использование ИМ в инженерных изысканиях обуславливает модернизацию компонентов системы, в связи с чем представленная модель претерпевает ряд изменений.

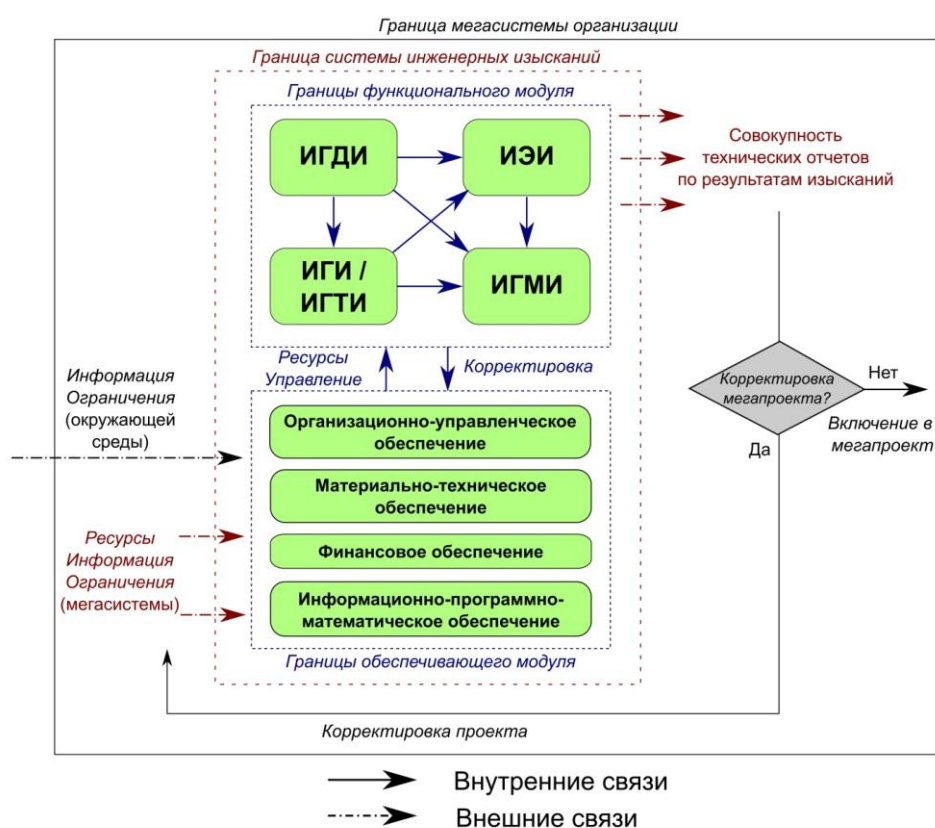


Рис. 2. Комплексная модель инженерных изысканий

Fig. 2. Complex model of engineering survey

Обновленная модель функционирования системы инженерных изысканий должна позволять получать в качестве выходных результатов технические отчеты, графические данные и трехмерную сводную модель инженерных изысканий. При этом очевидно, что входные параметры должны быть дополнены специальными требованиями к ИМ (дополнительные ограничения) и перечнем необходимого для работы программного обеспечения (дополнительные ресурсы).

Состав и структура инженерных изысканий остаются прежними, так как ИМ не предполагает создание новых подсистем. При этом изменяется содержание связей в рамках функционального модуля системы: основной единицей, характеризующей их, становятся элементы ИМ – цифровая модель рельефа (производная ИГДИ), модели данных гидрометеорологических, геолого-геотехнических и экологических наблюдений (производные ИГМИ, ИГИ / ИГТИ и ИЭИ соответственно).

Стоит отметить, что осуществление менеджмента ИМ (формирование сводной информационной модели инженерных изысканий и контроль создания ее частей) рассматривается как внешнее воздействие на систему со стороны мегасистемы.

Инженерно-экологические изыскания в контексте комплексной модели инженерных изысканий и внедрения технологий ИМ

Используя принцип иерархичности систем, рассмотрим подсистему ИЭИ в качестве отдельной системы. В таком случае для системы инженерно-экологических изысканий (далее подсистемы) рассмотренная выше система инженерных изысканий выступает как один из факторов внешней среды.

К входным параметрам для подсистемы инженерно-экологических изысканий относятся входы системы инженерных изысканий. Кроме того, на нее воздействуют другие подсистемы функционального модуля, предоставляющие сведения о рельефе, геологических и гидрологических особенностях участка изысканий. Результат функционирования подсистемы ИЭИ приводит к появлению в мегапроекте сведений об инженерно-экологическом состоянии территории проектирования в виде технического отчета.

В составе технического отчета⁵ по результатам инженерно-экологических изысканий выделяют пояснительную записку и графические материалы (табл. 2). Экологическая информация, содержащаяся в пояснительной записке, включает в себя оценку геоэкологического состояния территории, описание особенностей почвы и растительного покрова, состава фауны, а также вероятности возникновения опасных экзогенных геологических процессов и гидрологических явлений. Характеристика современного и прогнозируемого экологического состояния территории основана на зонировании территории по принципу наличия участков ограничений хозяйственной деятельности, загрязнения компонентов среды и предполагаемой деградации почвенного и растительного покрова в процессе дальнейших работ на исследуемом участке. Графическая часть визуализирует представленную в пояснительной записке информацию.

Процесс создания технического отчета по результатам ИЭИ представлен в виде схемы формирования экологического обеспечения проекта (рис. 3). Пояснительная записка содержит информацию, необходимую для атрибутирования и описания графических материалов, а графические материалы дополняют пояснительную записку данными о пространственном расположении объекта и описываемых в тексте объектов, процессов и явлений.

⁵ СП 47.13330.2016 «Инженерные изыскания для строительства. Основные положения»; СП 11-102-97 «Инженерно-экологические изыскания для строительства».

Таблица 2

Состав инженерно-экологического обеспечения проекта

Table 2

The composition of the engineering and environmental support of the project

Элемент	Описание
Пояснительная записка	Текстовая часть отчета по результатам ИЭИ, содержащая экологическую информацию в соответствии с нормативно-правовыми актами, регулирующими состав и содержание отчета
Графические материалы	Картосхемы, выполненные в двумерном виде с использованием ГИС, дополняющие и раскрывающие пояснительную записку

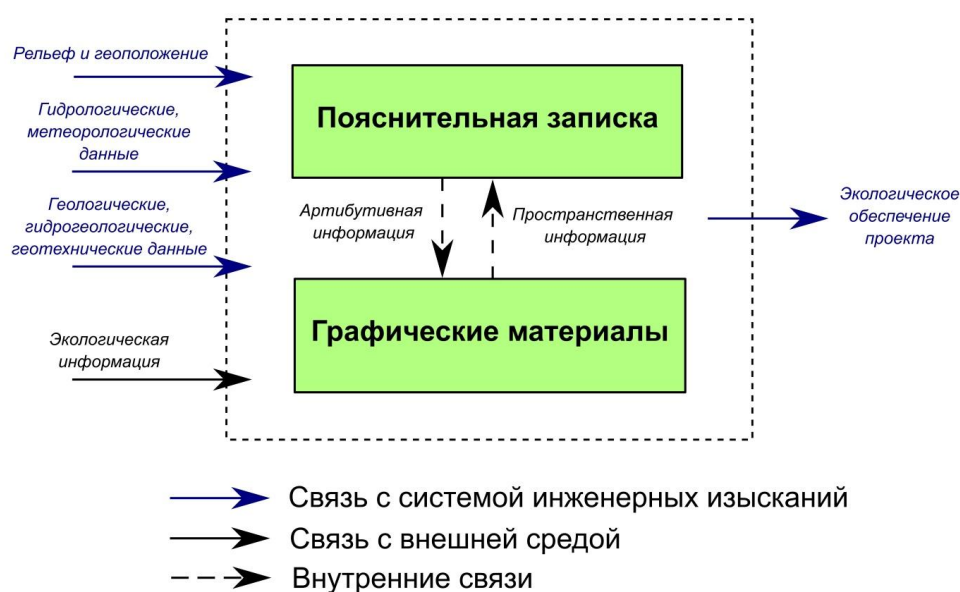


Рис. 3. Схема формирования экологического обеспечения проекта
 Fig. 3. Scheme for the formation of the environmental support of the project

Внедрение ИМ в процесс ИЭИ приводит к изменению выходных параметров подсистемы – в составе инженерно-экологического обеспечения проекта выделяется модель данных экологических наблюдений как часть сводной цифровой модели инженерных изысканий. В соответствии с этим трансформируются входные и выходные параметры подсистемы аналогично рассмотренным ранее входам и выходам системы инженерных изысканий.

Рассмотрено несколько вариантов состава экологического обоснования проекта в случае внедрения технологии ИМ. В первом варианте модель данных экологических наблюдений разрабатывается на основе графических материалов и выступает в качестве дополнительной составляющей экологического обеспечения проекта (рис. 4, а). Второй вариант предусматривает формирование модели данных экологических наблюдений как основы для создания графических материалов в традиционном (двумерном) и трехмерном видах (рис. 4, б).

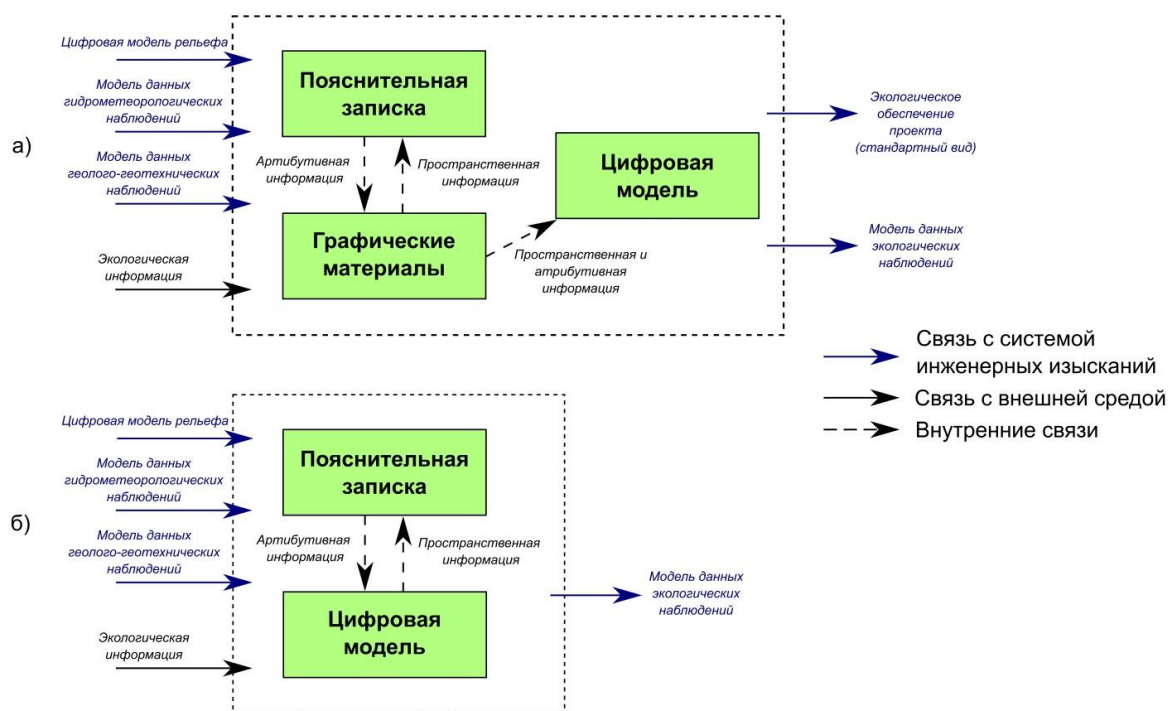


Рис. 4. Схема формирования инженерно-экологического обеспечения проекта с учетом внедрения технологии ИМ (а – первый вариант, б – второй вариант)

Fig. 4. Scheme of the formation of engineering and environmental support of the project, taking into account the introduction of BIM (a – first option, b – second option)

Построение и анализ плана-сценария развития системы инженерно-экологических изысканий с учетом внедрения ИМ

Применение системного анализа позволило выявить особенности системы инженерных изысканий и подсистемы инженерно-экологических изысканий, охарактеризовать их изменения при внедрении технологии ИМ в процесс работы. Эффективный переход от исходного к желаемому состоянию систем необходимо выполнять с учетом оценки рисков, имеющихся ресурсов и возможностей. Решение этой задачи возможно благодаря использованию комплекса методов форсайт-исследования [3] и риск-менеджмента [4]. Идея заключается в построении плана-сценария внедрения технологии ИМ с определением узловых точек – этапов, на которых необходимо осуществить выбор между альтернативными вариантами развития на основе оценки рисков, возможностей и имеющихся ресурсов.

Узловые точки и направления развития в плане-сценарии внедрения технологий ИМ в ИЭИ (рис. 5) выделены на основе анализа нормативно-правовой базы в сфере инженерных изысканий и ИМ, а также зарубежного и отечественного опыта информационного моделирования.

В рамках исследования выделено две узловые точки внедрения ИМ в процесс ИЭИ: выбор итогового состава экологического обеспечения проекта и выбор программного обеспечения для его достижения. При этом в зависимости от принятия решений формируется три варианта итогового состояния системы – первые два основаны на модернизации системы, а третий представляет собой возврат к исходному состоянию.

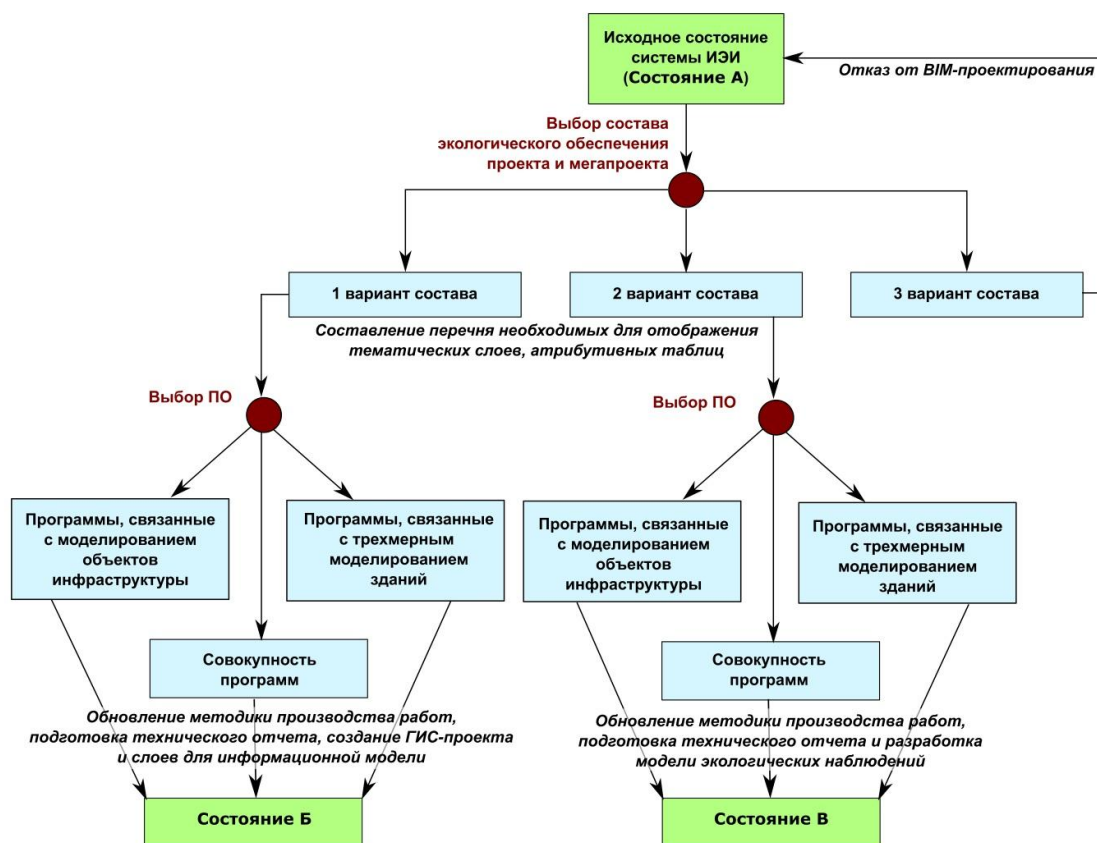


Рис. 5. План-сценарий внедрения технологии информационного моделирования в инженерно-экологические изыскания

Fig. 5. Scenario plan for the introduction of BIM in engineering and environmental surveys

В оценку рисков и возможностей в ходе принятия решений по выбору приемлемого для воплощения варианта (табл. 3) входит характеристика их содержания, вероятности возникновения, возможных путей управления ими, а также наиболее уязвимый из выделенных путей развития системы.

Наиболее уязвимым в узловой точке 1 является третий путь, подразумевающий возврат к исходному состоянию А. Данное заключение связано с предполагаемым изменением законодательства в пользу BIM-проектирования, а также с маловероятными положительными изменениями внутри системы в случае отказа от технологий ИМ.

Второй путь развития, ведущий к состоянию В, связан с высоким риском увеличения времени проектирования и возможности получения отрицательного заключения экспертов, а также общей массивностью при прочих равных факторах.

Преимуществом первого пути внедрения технологии ИМ является возможность возврата к исходному состоянию системы А или планомерного развития до состояния В. Таким образом, следует принять решение по выбору первого пути внедрения технологии ИМ как наиболее приемлемого из представленных.

Таблица 3

Оценка рисков и возможностей при выборе направления развития

Table 3

Assessment of risks and opportunities when choosing the direction of development

Содержание риска / возможности	Вероятность (по вариантам)	Пути управления	Уязвимый вариант
Узловая точка 1: выбор желаемого комплекса моделей			
Р и с к и	Получение отрицательного заключения Главгос-экспертизы РФ	1. Разработка документации в двух видах: стандарт-ный технический отчет и инженерная цифровая модель местности. 2. Обратная связь с экспертами в рамках ИМ	2
	Изменение законодательства в области ИМ	1. Пилотное проектирование. 2. Подготовка кадров. 3. Обратная связь со специалистами, разрабатывающими нормативную документацию	3
	Неверная оценка стоимости работ	1. Пилотное проектирование. 2. Подготовка кадров. 3. Изучение опыта ИМ	1, 2
	Рост времени работы над проектами		2
	Рост объемов выполненных работ	1. Пилотное проектирование. 2. Подготовка кадров	3
	Рост качества выпускаемой продукции		
Возможности	Рост количества заказов	1. Обратная связь с заказчиками. 2. Повышение доли BIM-проектов	

Содержание риска / возможности		Вероятность (по вариантам)	Пути управления	Уязвимый вариант
Узловая точка 2: выбор программного обеспечения				
Риски	Необходимость предоставления модели в другом ПО	Все – средняя	1. Мониторинг рынка заказов. 2. Использование при разработке ИМ формата IFC с открытой спецификацией	Все
	Отсутствие достаточного количества лицензий ПО	Все – средняя	1. Определение количества сотрудников, задействованных в разработке ИМ. 2. Выделение резервов на закупку ПО	Все
	Низкая совместимость BIM- и ГИС-систем	1.1, 2.1 – низкая 1.2, 1.3, 2.2, 2.3 – средняя	1. Обратная связь с разработчиками программ. 2. Анализ альтернативных вариантов программного обеспечения	1.2, 1.3, 2.2, 2.3
	Массивность файлов и сложность файловой структуры	1.1, 1.2, 1.3 – средняя 2.1, 2.2, 2.3 – высокая	1. Создание облачных хранилищ и серверных дисков для хранения. 2. Разделение BIM-проектов на этапы	2.1, 2.2, 2.3

Из трех вариантов используемого программного обеспечения наиболее эффективным представляется работа с набором программ, в этом случае следует необходимая оценка возможностей по оснащению системы необходимым программным обеспечением. На наш взгляд, наиболее выгодным как с точки зрения оценки рисков и возможностей, так и с точки зрения использования ресурсов организации становится выбор пути 1.1. Он предполагает формирование пояснительной записки, графических материалов и модели данных экологических наблюдений с использованием программного обеспечения, основанного на проектировании инфраструктуры в рамках ИМ.

Выводы

В ходе исследования разработана методика перехода к ИМ в рамках инженерных изысканий на примере блока инженерно-экологических изысканий, начальное и конечные состояния функционирования которого определены и описаны с помощью методов системного анализа. Пути внедрения ИМ установлены благодаря форсайт-исследованию, а риск-менеджмент позволил обосновать выбор наиболее эффективного направления.

В методике выделены следующие ключевые этапы:

- 1) построение комплексной модели текущего состояния системы изысканий;
- 2) характеристика изменений в системе при внедрении технологии ИМ;
- 3) создание плана-сценария внедрения технологии ИМ с выделением альтернативных путей развития системы;
- 4) выбор наиболее эффективного варианта развития системы с помощью оценки рисков, возможностей и доступных ресурсов в узловых точках развития.

В итоге при отсутствии дополнительных ограничений и в условиях зарождения нормативной правовой базы ИМ в России выглядит наиболее эффективным представление результатов изысканий в виде стандартных технических отчетов, дополненных цифровыми моделями по видам изысканий, с постепенным усилением роли информационного моделирования. При этом рекомендуется использовать программное обеспечение трехмерного моделирования инфраструктуры.

Список литературы

1. **Sacks R., Eastman Ch., Lee G., Teicholz P.** BIM handbook: a guide to building information modeling for owners, managers, designers, engineers and contractors. 3rd ed. USA: Jhon Wiely and Sons, Inc., 2018, 659 p. DOI 10.1002/9781119287568
2. **Иванова Т. Ю.** Теория организации: Учебник. 4-е изд. М.: КНОРУС, 2012. 432 с.
3. **Кинэн М.** Технологический Форсайт: Международный опыт // Форсайт. 2009. Т. 3, № 3. С. 60–68. DOI 10.17323/1995-459x.2009.3.60.68
4. **Капустина Н. В.** Теоретико-методологические подходы риск-менеджмента: Монография. М.: ИНФРА-М, 2016. 140 с.

References

1. **Sacks R., Eastman Ch., Lee G., Teicholz P.** BIM handbook: a guide to building information modeling for owners, managers, designers, engineers and contractors. 3rd ed. USA: Jhon Wiely and Sons, Inc., 2018, 659 p. DOI 10.1002/9781119287568
2. **Ivanova T. Yu.** Teoriya organizatsii [Organization theory] Textbook. 4th ed. Moscow, KNORUS, 2012, 432 p (in Russ.)
3. **Keenan M.** Tekhnologicheskii Forsayt: mezhdunarodnyy opyt [Technology Foresight: International Experience]. *Foresight-Russia*, 2009, vol. 3, no. 3, pp. 60–68. (in Russ.). DOI 10.17323/1995-459x.2009.3.60.68

4. **Kapustina N. V.** Teoretiko-metodologicheskie podhody risk-menedzhmenta [Theoretical and methodological approaches to risk management]. Monograph. Moscow, INFRA-M, 2016, 140 p. (in Russ.)

Информация об авторах

Антон Сергеевич Флеенко, магистрант второго года обучения
Александр Федорович Демьяненко, доктор технических наук, профессор
Author ID 310010

Information about the Authors

Anton S. Fleenko, 2nd year Master Student
Alexander F. Demianenko, Doctor of Sciences (Techniques), Professor
Author ID 310010

*Статья поступила в редакцию 10.06.2021;
одобрена после рецензирования 01.07.2021; принята к публикации 01.07.2021
The article was submitted 10.06.2021;
approved after reviewing 01.07.2021; accepted for publication 01.07.2021*

Правила оформления текста рукописи

Авторы представляют статьи на русском или английском языке объемом от 0,5 авторского листа (20 тыс. знаков) до 1 авторского листа (40 тыс. знаков), включая иллюстрации (1 иллюстрация форматом 190 × 270 мм = 1/6 авторского листа, или 6,7 тыс. знаков). Публикации, превышающие указанный объем, допускаются к рассмотрению только после индивидуального согласования с редакцией журнала.

Текст рукописи должен быть представлен в редколлегию в виде файла MS Word (.doc, .docx). Гарнитура Times New Roman, размер шрифта 11, межстрочный интервал 1, размеры полей – стандартные значения текстового редактора. Форматирование – выравнивание по ширине страницы, переносы слов включены, каждый новый абзац начинается с красной строки. Не допускается ручное форматирование абзацев (пробелами, лишними переводами строк, разрывами страниц).

Структура статьи

- Индекс УДК (универсальной десятичной классификации). Выравнивание по левому краю
- Название статьи. Выравнивание по центру, полужирный шрифт
- ФИО авторов (полностью). Выравнивание по центру, полужирный шрифт
- Места работы всех авторов. Выравнивание по центру, курсив
- Адреса электронной почты, ORCID авторов
- Аннотация статьи
- Ключевые слова, не более 10
- Благодарности, сведения о финансовой поддержке
- Название статьи **на английском языке**. Выравнивание по центру, полужирный шрифт
- ФИО авторов **на английском языке** (полностью). Выравнивание по центру, полужирный шрифт
- Места работы авторов **на английском языке**. Выравнивание по центру, курсив
- Аннотация статьи **на английском языке (Abstract)**, 200–250 слов
- Ключевые слова **на английском языке (Keywords)**, не более 10
- Благодарности, сведения о финансовой поддержке **на английском языке**, если есть соответствующий раздел на русском языке (**Acknowledgements**)
- Основной текст
- Список литературы / **References**
- Сведения об авторах

Требования к оформлению основного текста и иллюстративных материалов

Основной текст должен быть представлен в структурированном виде, рекомендуется использовать подзаголовки – например: Введение, Методика..., Выводы, Результаты, Заключение.

Подзаголовки отделяются и набираются полужирным шрифтом. В целях выделения частей текста и отдельных слов и словосочетаний допускается использование курсива или полужирного шрифта. Подчеркивание, разрядка, изменение основного кегля и выделение цветом не используются.

Иллюстрации к рукописи статьи должны быть приложены в виде отдельных файлов. При этом в тексте должно содержаться включенное изображение с указанием имени файла. Все иллюстрации, содержащие схемы, графики, алгоритмы и т. п., должны быть представлены в векторном виде (.ai, .eps, .cdr). Скриншоты и другие растровые изображения должны быть представлены в максимально высоком качестве, без каких-либо потерь и искажений (.jpg, .tif). Все иллюстрации должны иметь подрисовочную подпись – свое название. Надписи к таблицам и подписи к иллюстрациям приводятся **на двух языках (русском и английском).**

Примеры:

Рис. 1. Диаграмма производительности...

Fig. 1. Performance diagram...

Таблица 1

Сравнение алгоритмов...

Table 1

Comparison of algorithms...

Нумерация последовательная и неразрывная от начала статьи. Не допускается использование других наименований, кроме «Рис.» / «Fig.», «Таблица» / «Table», и усложнение нумерации (например, «Рис. 3.2.»). Ссылка на иллюстрацию в тексте должна быть приведена в круглых скобках, например: (рис. 1), (табл. 1).

Формулы должны быть набраны с использованием редактора MathType либо встроенного редактора формул MS Word. Кегль основных символов – 11, греческие символы набираются прямым шрифтом, латинские – курсивом. Нумеруются только те формулы, на которые автор ссылается в тексте.

Abstract

Аннотация статьи на английском языке (Abstract) не должна быть дословным переводом русскоязычной аннотации. Раздел Abstract, как и основной текст, должен быть структурирован, в нем должно содержаться описание цели работы, методов исследования, научной значимости, выводов / результатов. Требуется качественный перевод на английский язык (при необходимости просим авторов обращаться к профессиональным переводчикам). **Объем Abstract 200–250 слов.**

Список литературы / References

Список литературы и список литературы на английском языке (References) размещаются в общем разделе. Рекомендуемое количество цитируемых в статье источников – не менее 10, в список желательно включать ссылки на актуальные работы по теме исследования, особенно в иностранных периодических изданиях.

В тексте статьи ссылки на литературу указываются цифрами в квадратных скобках, при необходимости указываются номера страниц, например: [2; 3. С. 15].

Список литературы нумеруется в порядке цитирования и оформляется в соответствии с ГОСТ Р 7.0.5-2008 на библиографическое описание (знаки тире в описании опускаются). Ссылки на неопубликованные работы, а также на Интернет-ресурсы (кроме электронных изданий, поддающихся библиографическому описанию) оформляются в виде сноски.

В Список литературы ссылки на источники следует включать на оригинальном языке опубликования. Каждый источник должен быть также оформлен на английском языке (References) по международному стандарту для публикаций в области информатики IEEE Style со следующими отличиями:

- инициалы авторов указываются после фамилии;
- название статьи не берется в кавычки, отделяется точкой;
- отсутствует союз «and» перед фамилией последнего автора;
- в диапазоне страниц – удвоенная «р» (например, «pp. 2–9»);

• год издания указывается после места издания (для книг) и сразу после названия журнала (для периодики).

Перевод источника на английский язык:

• если источник имеет выходные данные на английском языке, то для формирования References **следует использовать именно эти данные**;

• если оригинальная публикация не содержит выходных данных на английском языке, то допускается транслитерация названия материала на латинский алфавит в сочетании с переводом на английский язык в квадратных скобках. В конце описания указывается, на каком языке написана эта работа, например, (in Russ.). При транслитерации можно воспользоваться Интернет-ресурсом <http://ru.translit.ru/>, рекомендуется выбрать стандарт BSI. Место издания не транслитерируется, указывается полностью на английском языке, например: Moscow. Название издательства / издателя, как правило, транслитерируется. Для журналов, у которых есть официальное название на английском языке, – использовать его (проверить на сайте журнала, или, например, в библиотеке WorldCat), если названия на английском языке нет, использовать транслитерацию по системе BSI. Не следует самостоятельно переводить названия журналов.

Если у цитируемого источника есть **цифровой идентификатор DOI** (<https://search.crossref.org>), его требуется обязательно указывать в конце библиографической ссылки.

Примеры оформления ссылок. Каждый источник в том же пункте дублируется на английском языке (References).

Источник на русском языке, перевод на английский доступен в метаданных статьи

1. Журавлев С. С., Рудометов С. В., Окольников В. В., Шакиров С. Р. Применение модельно-ориентированного проектирования к созданию АСУ ТП опасных промышленных объектов // Вестник НГУ. Серия: Информационные технологии. 2018. Т. 16, № 4. С. 56–67. DOI 10.25205/1818-7900-2018-16-4-56-67

Zhuravlev S. S., Rudometov S. V., Okolnishnikov V. V., Shakirov S. R. Model-Based Design Approach for Development Process Control Systems of Hazardous Industrial Facilities. *Vestnik NSU. Series: Information Technologies*, 2018, vol. 16, no. 4, pp. 56–67. (in Russ.) DOI 10.25205/1818-7900-2018-16-4-56-67

Источник на английском языке. Оформляем согласно требованиям для References. Приводим только 1 раз.

2. Telnov V. I. Optimization of the Beam Crossing Angle at the ILC for E + e- and yy Collisions. *Journal of Instrumentation*, 2018, vol. 13, no. 03, pp. P03020–P03020. DOI 10.1088/1748-0221/13/03/p03020

Метаданные источника доступны только на русском языке

3. Жижимов О. Л., Федотов А. М., Шокин Ю. И. Технологическая платформа массовой интеграции гетерогенных данных // Вестник НГУ. Серия: Информационные технологии. 2013. Т. 11, вып. 1. С. 24–41.

Zhizhimov O. L., Fedotov A. M., Shokin Yu. I. Tekhnologicheskaya platforma massovoi integratsii geterogennykh dannyykh [Technology Platform for the Mass Integration of Heterogeneous Data]. *Vestnik NSU. Series: Information Technologies*, 2013, vol. 11, no. 1, pp. 24–41. (in Russ.)

Сведения об авторах

Последний раздел статьи – информация об авторе / авторах **на русском и английском языках**:

- ФИО полностью, ученая степень, ученое звание;
- идентификаторы автора, такие как ResearcherID (всем авторам рекомендуется использовать данные сервисы для ведения актуального списка своих публикаций);

- контактный телефон (не публикуется).

Если статья представляется на английском языке, необходимо приложить перевод на русский язык названия, аннотации, ключевых слов, сведений об авторе.

Доставка материалов

Материалы предоставляются в редакцию по электронной почте inftech@vestnik.nsu.ru.

Порядок рецензирования

Все статьи сначала проходят проверку на заимствование и только после этого отправляются на рецензирование. Редакционный совет не допускает к публикации материал, если имеется достаточно оснований полагать, что он является плагиатом.

Тип рецензирования статей – двухуровневое, одностороннее анонимное («слепое»).

Для каждой статьи редколлегией выбираются рецензенты, научная деятельность которых связана с темой представленного материала. Ответственный секретарь журнала обращается к ним с просьбой дать экспертную оценку статье либо помочь организовать рецензирование.

Рецензии для журнала «Вестник НГУ. Серия: Информационные технологии» составляются по единой схеме и подразумевают оценку по следующим критериям: соответствие тематике журнала, оригинальность и значимость результатов, качество изложения материала.

Заполненный бланк рецензии высылается на электронный адрес редакции. В зависимости от экспертных заключений статья может быть принята редакционным советом к опубликованию, рекомендована автору к доработке (с последующим повторным рецензированием либо без него) или отклонена (с предоставлением автору мотивированного отказа). Автору на электронный адрес высылается текст рецензии без указания ФИО рецензента и его контактных данных.

Все рецензии хранятся в редакции журнала не менее 5 лет. Редколлегия журнала обязуется при поступлении соответствующего запроса направлять копии рецензий в Министерство науки и высшего образования Российской Федерации.