

А.В. Козловский, Э.В. Мельник, А.Н. Самойлов

**ПРИМЕНЕНИЕ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ ПРИ ОРГАНИЗАЦИИ
ПОПОЛНЕНИЯ БАЗЫ ЗНАНИЙ О МЕТОДАХ ОБРАБОТКИ ИНФОРМАЦИИ
В СИСТЕМАХ ПРИКЛАДНОЙ ФОТОГРАММЕТРИИ**

Рассмотрены вопросы, связанные с автоматизацией процедуры синтеза систем прикладной фотограмметрии. Такие системы служат для измерения и учета объектов по изображениям и в настоящее время широко применяются в различных областях деятельности, таких как картографирование, археология и аэрофотосъемка. Широкому применению также способствует повышение доступности и мобильности устройств для получения изображений. Все это обусловило проведение активных исследований, направленных на разработку методического обеспечения для систем прикладной фотограмметрии. Отслеживание в ручном режиме появления новых методов и алгоритмов фотограмметрической обработки информации для широкой номенклатуры областей применения достаточно затруднительно, что делает актуальной автоматизацию данной процедуры. Предлагаемое в статье решение основано на использовании базы знаний о методах обработки информации в системах прикладной фотограмметрии, основными элементами которой являются нечеткая онтология предметной области и база данных, что логично, т.к. информация о предметной области может быть достаточно легко структурирована. В качестве основы для онтологии было взято существующее решение, которое было дополнено на основе результатов анализа текущего состояния предметной области. Полученная онтология далее использована для поиска и классификации методов обработки информации в системах прикладной фотограмметрии и заполнения базы знаний. В связи с активизацией разработки новых методов обработки информации в системах прикладной фотограмметрии возникает необходимость модификации онтологии и пополнения базы данных, т. е. пополнения базы знаний. Важным источником информации для этого является Интернет. Для автоматизации поиска данных о методах обработки информации и пополнения базы знаний целесообразно использовать большие языковые модели, благодаря которым упрощается решение нескольких задач в области обработки естественного языка, которые включают кластеризацию и формирование новых сущностей для классификации. Соответствующий метод описан в статье. Для метода приведены результаты тестирования его работоспособности. В рамках решения задач проведен сравнительный анализ больших языковых моделей, в результате которого была выбрана модель RoBERTa.

Машинное обучение; графы; обработка изображений; алгоритмы классификации; искусственный интеллект; системы поддержки принятия решений; автоматизация процессов; большие языковые модели; онтология.

A.V. Kozlovskiy, E.V. Melnik, A.N. Samoylov

**A LARGE LANGUAGE MODELS APPLICATION IN ORGANIZATION
OF REPLENISHMENT OF THE KNOWLEDGE BASE ON METHODS
OF INFORMATION PROCESSING IN SYSTEMS OF APPLIED
PHOTOGRAMMETRY**

The article deals with the issues related to the automation of the procedure of synthesis of applied photogrammetry systems. Such systems serve to measure and account for objects from images and are now widely utilized in various fields of activity, such as mapping, archaeology and aerial photography. Increasing availability and mobility of imaging devices has also contributed to the widespread application. All this has led to active research aimed at developing methods and algorithms for applied photogrammetry systems. Manual tracking of new methods and algorithms of photogrammetric information processing for a wide range of application areas is quite difficult, which makes the automation of this procedure urgent. The solution proposed in the article is based on the use of a knowledge base of information processing methods in applied photogrammetry systems, the main elements of which are a fuzzy ontology of the subject area and a database, which is logical, since the information about the subject area can be structured quite easily. As a basis for the ontology, an existing solution was taken, which was supplemented based on the results of analyzing the current state of the subject area. The resulting ontology was further used to search and classify information processing methods in applied photogrammetry systems and to populate the knowledge base. Due to the intensification of the development of new methods of

information processing in the systems of applied photogrammetry, there is a need to modify the ontology and to replenish the database, i.e. to replenish the knowledge base. The Internet is an important source of information for this purpose. To automate the search for data on information processing methods and ontology modification, it is reasonable to use large language models. To automate data mining of information processing methods and to populate the knowledge base, it is useful to use large language models that simplify several natural language processing tasks, which include clustering and formation of new entities for classification. The corresponding method is described in the paper. For the method the results of testing its performance are given. As part of problem solving, a comparative analysis of large language models has been carried out, resulting in the RoBERTa model.

Machine learning; graphs; image processing; classification algorithms; artificial intelligence; decision support systems; process automation; large language models; ontology.

Введение. В настоящее время существует множество задач из различных областей, где возможно применение современных средств автоматизации для ускорения и удешевления производственных процессов. Такие средства позволяют заменить специально подготовленный персонал. Одной из важных областей, для которой актуально создание средств автоматизации, является фотограмметрия [1, 2]. Фотограмметрия ориентирована на решение проблем измерения (площадь, объём [3], местоположение [4, 5]) и подсчёта объектов на изображениях. Для этого применяются различные методы обработки графических данных, например оптическая коррекция изображений [6], объединение нескольких кадров в единый снимок, 3D-моделирование различных объектов [7–9]. В связи с актуальностью решения задач прикладной фотограмметрии исследования в данной области характеризуются высокой интенсивностью, что обуславливает достаточно частое появление новых и модифицированных методов, предназначенных для повышения эффективности решения указанных задач. Из-за этого ручное отслеживание обновлений и изменений становится затруднительным. В данном случае автоматизация может быть реализована на основе методов машинного обучения: может быть создано программное обеспечение, периодически выполняющее поиск в открытых источниках новых и модифицированных методов и добавляющее их в собственную базу знаний. Это позволяет сократить время и ресурсы, необходимые для проведения исследований, а также повысить точность и объективность получаемых результатов.

Несмотря на то, что автоматизация процессов является ключевым направлением развития во многих областях, подобные решения не получили широкого распространения зачастую из-за ограниченных возможностей структурирования данных вследствие особенностей предметной области: в отличие от фотограмметрии, такие области как медицина, право и другие, имеют гораздо более сложные и неопределённые факторы, влияющие на выбор того или иного метода.

Отсутствие достаточного количества структурированных данных также является препятствием для автоматизации: для построения эффективной базы знаний требуется большой объём структурированной информации о методах, условиях их применения и результатах. Также следует отметить имевший место до недавнего времени недостаточный уровень развития технологий обработки естественного языка и машинного обучения для автоматического анализа и структурирования информации из различных источников.

Существующие решения и необходимость их модификации. В ходе исследования предметной области были рассмотрены работы, в которых предложены решения, позволяющие реализовать автоматизированное обновление базы знаний о методах решения задач в области прикладной фотограмметрии [13, 14]. В данных работах для описания предметной области предложено использовать нечёткую онтологию, поскольку невозможно обеспечить точное совпадение значений параметров в базе знаний с фактически имеющимися. Данная онтология является частью базы знаний, её структура приведена в [10].

Существующая онтология требует корректировки. Одной из причин является, например, расширение номенклатуры устройств, которые целесообразно применять при решении задач фотограмметрии. Чтобы получать качественные изображения, изначально использовались профессиональные устройства, обладающие соответствующими характе-

ридиками при ограничениях по мобильности и стоимости. В настоящее время стали применяться мобильные устройства, ориентированные, в первую очередь, на получение изображений, пригодных для простого просмотра пользователем, и для достижения требуемого, с точки зрения решения задач фотограмметрии, результата, например, в условиях недостаточной освещенности, при использовании таких устройств требуется постобработка. При этом номенклатура методов обработки изображений достаточно широка (и постоянно дополняется) и в разных случаях используются разные их сочетания. Таким образом, разработанную ранее онтологию следует модифицировать с учетом расширения номенклатуры применяемых устройств и соответствующего методического обеспечения.

Необходимость модификации онтологии также связана с тем, что фотограмметрия имеет гораздо больше применений, чем подразумевает существующая структура данных. Изменения существующей структуры будут носить комплексный характер. Это должно предоставить большую гибкость, которая позволит динамически дополнять базу знаний.

Необходимые изменения. Первой и самой важной модификацией станет добавление класса «Задача». Он будет строго задавать желаемый конечный результат, который должен быть представлен после работы подобранных методов. В данном классе будут подклассы, задающие общие свойства для групп задач. Например, одним из подклассов станут измерительные задачи. Для них будет обязательным указанием того, какую именно величину в данный момент необходимо определить. То есть компоненты класса «Измерение» будут перенесены в подкласс измерительных задач. Однако ряд возможных измерений тоже будет дополнен понятием «Координаты». Оно будет полезно при определении местоположения объекта. Также заранее можно определить такие подклассы для «Задачи», как «3D-моделирование», «Склейка / Повышение чёткости», «Коррекция», «Прогнозирование / Проектирование» и «Калибровка». Под последним элементом имеется в виду ручное занесение всех параметров. Он будет неразрывно связан с методом «Калибровка на ходу». Перечисленные ранее подклассы тоже будут разделяться. «Коррекция» будет делиться на подклассы «Оптическая» и «Цветовая», «Прогнозирование / Проектирование» – на «Природное» и «Урбанистическое», «3D-моделирование» – на «Объект» и «Поверхность / Рельеф».

Помимо этого, методы прикладной фотограмметрии можно разделить на основные и вспомогательные. Так, к основным будут отнесены те, что проводят анализ изображений и дают оценку заданных параметров; к вспомогательным же будут отнесены те, что либо подготавливают само изображение для получения впоследствии более точных результатов, либо формируют их в форме удобной для восприятия человека. Однако необходимо учесть, что основная задача может быть именно в корректировке изображения, поэтому некоторые вспомогательные методы могут выступать и в роли основных.

Также стоит учитывать, что различные методы могут решать одни и те же или схожие задачи. Это свойство можно использовать для неоднозначных и сложных условий съёмки. Например, можно будет применять два и более методов одновременно, или «параллельно», объединяя их результат при помощи алгоритмов фильтрации и коэффициентов доверия. Для учёта этих изменений необходимо будет добавить к вспомогательным методам алгоритмы фильтрации полученных результатов. Помимо таких сочетаний методов возможно будет также и «последовательное» использование. Это будет справедливо как для цепочки «вспомогательный – основной» (или наоборот), так и для «основной – основной». Под понятием «основной» здесь может пониматься и ранее приведённый вариант «параллельного» применения. Стоит отметить, что возможность использовать метод «параллельно» или «последовательно» будет отмечаться непосредственно в онтологии. Таким образом, будет обеспечена возможность выстраивать любые последовательности применения методов для создания уникальных алгоритмов для различных условий, распознанных системой.

Изменения, подобные первым упомянутым, будут также применены к классу «Ситуация». Данный класс будет включать степень освещённости, тип источника света, время суток, погодные условия. Этот перечень атрибутов должен исчерпывающе описывать условия съёмки.

Следующей модификацией будет добавление класса, описывающего настройки фотоаппаратуры. В данный класс войдут такие параметры камеры, как диафрагма, выдержка, баланс белого и светочувствительность. Это необходимо, для того чтобы иметь более подробные сведения для более точного результата, так как такие данные позволят подбирать методы коррекции изображений. Также эта информация поможет изначально задавать настройки так, чтобы минимизировать потребность в самой коррекции. К тому же параметры камеры смогут помочь оценить, возможна ли съёмка в принципе при имеющемся оборудовании.

Последним изменением в структуре онтологии будет дополнение в класс CBR (Case Based Reasoning). Он должен стать состоять из четырёх элементов: задача, ситуация, параметры аппаратуры и как вывод из них метод или их связка (алгоритм). Онтология, построенная по предложенным принципам, должна обеспечить основу наподобие каркаса или скелета, которую будет легко масштабировать и дополнять.

Структура онтологии с предложенными модификациями изображена на рис. 1. Она точнее описывает связи между классами и позволяет производить более детальную классификацию. Однако для хранения больших объемов данных онтология не подходит. Для этой информации в дополнение необходимо отдельное хранилище, для реализации которого хорошо подойдёт документно-ориентированная база данных, в которой будут храниться текстовые поля с описанием методов и ссылкой на программную реализацию. Более подробно это будет рассмотрено в следующих пунктах.

Выбор языковой модели. Процесс интеграции новых методов фотограмметрии в базу знаний начинается с обработки текстовых описаний, собранных из различных источников. Для этого используются возможности больших языковых моделей (LLM) на этапе токенизации и генерации эмбедингов.

В ходе исследования проведен сравнительный анализ нескольких предобученных LLM [11–13], среди которых:

- ♦ **BERT (Bidirectional Encoder Representations from Transformers):** эффективен для анализа контекста слов и построения обобщенных представлений текста [14];
- ♦ **RoBERTa (A Robustly Optimized BERT Pretraining Approach):** улучшенная версия BERT, обученная на большем объеме данных и демонстрирующая более высокую точность в задачах понимания языка [15];
- ♦ **GPT-2/3 (Generative Pre-trained Transformer):** мощная модель, способная генерировать связный и грамматически правильный текст, а также хорошо улавливающая семантические связи между словами.

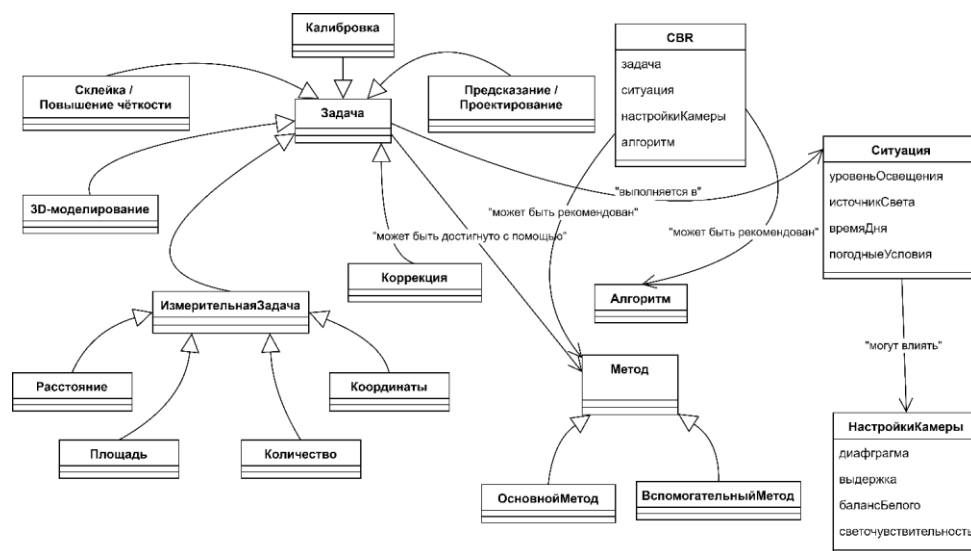


Рис. 1. Модели онтологий

В результате экспериментов с различными LLM и их конфигурациями была выбрана модель RoBERTa, которая показала наилучшие результаты в задачах классификации текстов по тематике фотограмметрии и извлечения ключевой информации о методах [16–18].

Токенизация. Применяется предварительно обученный токенизатор, связанный с моделью RoBERTa. Этот токенизатор разбивает текст на токены, учитывая специфику языка и обучающего набора данных RoBERTa. Важным аспектом является способность токенизатора корректно обрабатывать специфическую терминологию, характерную для области фотограмметрии. Токенизатор RoBERTa продемонстрировал хорошие результаты в распознавании и сохранении целостности таких терминов, как «Structure-from-Motion» (SfM), «Multi-View Stereo» (MVS), «Bundle Adjustment» и др. Токенизатор разбивает текст не только по пробелам, но и сохраняет целостность специфических терминов, представляя их как единый токен.

Для векторизации текста используются предобученные слои модели RoBERTa, которые уже кодируют семантическую информацию о словах и их контексте.

На вход модели RoBERTa подается последовательность токенов, полученная на предыдущем этапе. Модель анализирует эту последовательность, учитывая взаимосвязи между словами, и генерирует для каждого токена вектор фиксированной размерности – «эмбединг». Полученные «эмбединги» отражают не только значение отдельных слов, но и их смысл в контексте всего предложения или даже документа. Это позволяет улавливать более тонкие смысловые особенности в описаниях методов фотограмметрии [19, 20].

После преобразования текстовых описаний методов фотограмметрии в информационно-насыщенные «эмбединги» следующим этапом исследовательской работы стало извлечение ключевых характеристик каждого метода. Для этого использовалась способность LLM понимать естественный язык и отвечать на вопросы.

Разработка промптов является необходимым этапом к корректному извлечению информации, требующим тщательной формулировки запросов, предоставляемых LLM. ходе экспериментов с различными формулировками были найдены наиболее строгие и однозначные, исключающие двойную интерпретацию.

Примеры промптов для определения области применения включают:

- ◆ *«Опишите, для решения каких задач фотограмметрии применяется этот метод.»*
- ◆ *«В каких областях (археология, геодезия, др.) этот метод фотограмметрии нашел применение?»*

Для выявления требований к оборудованию:

- ◆ *«Какое оборудование (тип камеры, датчики, программное обеспечение) необходимо для реализации этого метода?»*
- ◆ *«Существуют ли какие-либо специфические требования к характеристикам оборудования (разрешение камеры, точность датчиков)?»*
- ◆ *Для определения связи с другими методами:*
- ◆ *«Какие методы фотограмметрии часто используются в сочетании с этим методом?»*
- ◆ *«Является ли этот метод модификацией или расширением какого-либо другого метода?»*

Включение дополнительного контекста, такого как название метода или краткое описание его сути, значительно повышает релевантность и точность ответов LLM.

Ответы LLM, хотя и формулируются на естественном языке, требуют дополнительной обработки для извлечения конкретной информации. Для этого используется комбинация методов, включая регулярные выражения для извлечения форматированных данных, таких как названия методов или типов оборудования, и методы NLP для анализа грамматической структуры предложений и выделения ключевых слов и фраз. Извлеченная информация преобразуется в структурированный формат с заполнением заранее определенных слотов в базе знаний.

После извлечения при помощи LLM эмбедингов методов фотограмметрии, наполненных информацией, следующим важным этапом исследования стала кластеризация. Целью кластеризации являлось объединение методов в группы на основе их семантической близости, что позволило выявить скрытые связи и закономерности в наборе данных.

Для измерения смысловой близости между эмбедингами методов был выбран метод косинусного расстояния. Анализ различных метрик сходства показал, что косинусное расстояние лучше всего подходит для задачи определения смысловой близости между описаниями методов. Преимущества косинусного расстояния:

- ◆ игнорирование длины вектора;
- ◆ чувствительность к смысловым связям.
- ◆ Для кластеризации были протестированы несколько алгоритмов:
- ◆ **k-means**: быстрый и простой алгоритм, но требующий задания числа кластеров заранее;
- ◆ **DBSCAN**: автоматически определяет число кластеров и выделяет выбросы, но чувствителен к выбору параметров;
- ◆ **иерархическая кластеризация**: позволяет визуализировать иерархию кластеров, но может быть вычислительно дорогостоящей для больших наборов данных.

В результате экспериментов наилучшие результаты показал алгоритм DBSCAN. Данный алгоритм обеспечил автоматическое определение оптимального числа кластеров, соответствующего интуитивному представлению о структуре предметной области.

Финальный этап исследования заключался в интеграции извлечённой и обработанной информации о методах фотограмметрии непосредственно в базу знаний. Этот процесс позволил представить знания в машиночитаемом формате, что открывает возможности для автоматического поиска, анализа и использования информации.

Первым шагом была генерация RDF triples. В качестве примера были выделены основные разделы для классификации, такие как «Метод», «ОбластьПрименения» и «Оборудование», а также отношения между ними, например, «используетсяДля» и «требует». На основе этого анализа были разработаны шаблоны для генерации RDF triples, каждый из которых соответствовал конкретному отношению в онтологии и определял, какие данные необходимо извлечь из текста и как их преобразовать в формат RDF. Пример шаблона может выглядеть так:

Шаблон: <метод> <область применения> <значение>.

Этот шаблон описывает отношение «область применения» между отдельным методом и конкретной областью, в которой он применяется. Заполнение шаблонов выполнялось с использованием LLM, которые служили интерпретаторами. LLM не только извлекали информацию, но и интерпретировали её в контексте онтологии, анализируя данные и подставляя их в соответствующие слоты шаблонов. Например, LLM могла определить, что метод «Structure-from-Motion» используется для «3D-моделирования зданий» и сформировать соответствующий RDF triple:

Заполненный шаблон: <Structure-from-Motion> <3D-моделирование зданий>.

На следующем этапе интеграции создавались новые сущности для каждого нового метода фотограмметрии, описанного в текстах. Эти сущности добавлялись в онтологию в качестве новых элементов класса «Метод». Сгенерированные RDF triples использовались для установления связей между новыми сущностями и существующими в онтологии классами. Визуализация собранной базы знаний в виде графа представлена на рис. 2.

Возникшие на этом этапе противоречия разрешаются с помощью набора эвристик, например, если в разных источниках для одного и того же метода указывались разные области применения. Эти эвристики основываются на приоритете источников, частоте упоминаний и других факторах, что позволило эффективно устранять конфликты.

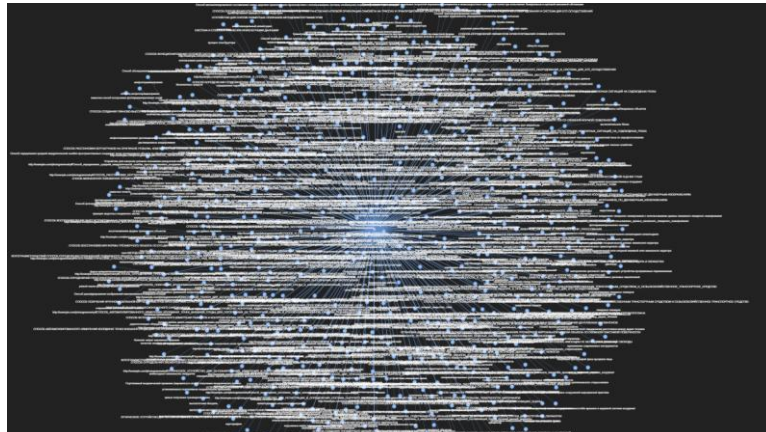


Рис. 2. Общий граф онтологии

Пример конкретного метода фотограмметрии и связанных подклассов представлены на рис. 3.

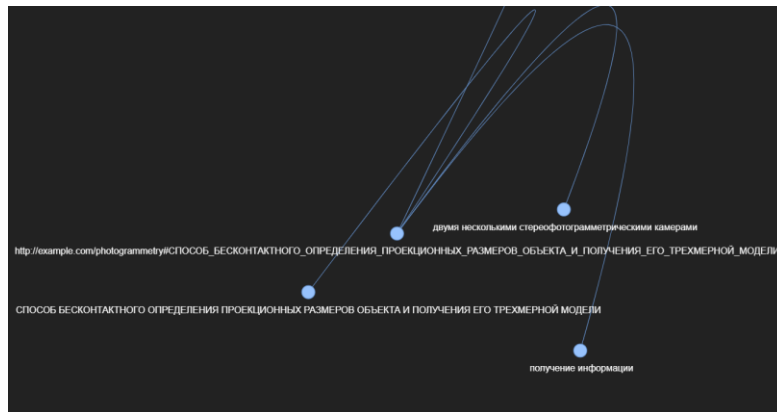


Рис. 3. Пример подграфа для метода фотограмметрии в онтологии

Хранение данных может осуществляться в формате OWL, который является стандартным для представления знаний в Семантическом вебе.

Заключение. Предложенная в данной работе модифицированная онтология для формирования и применения базы знаний для выбора методов решения задач прикладной фотограмметрии учитывает современный уровень развития методической базы и аппаратного обеспечения в области прикладной фотограмметрии. Это позволяет использовать ее в составе базы знаний, используемой при проектировании проблемно-ориентированных систем в данной предметной области.

Для интеграции новых методов в базу знаний предложено использовать большие языковые модели. Это позволяет автоматически базу знаний о современных методах, применяемых в системах прикладной фотограмметрии, что обеспечит возможность определения в ходе проектирования оптимальные сочетания методов для конкретной задачи фотограмметрии, учитывая особенности каждого конкретного случая.

Полученные результаты могут быть применены при синтезе систем прикладной фотограмметрии для различных областей деятельности, включая картографирование, археологию, аэрофотосъемку и другие.

Благодарности. Исследование проведено в рамках студенческого научного проекта N 4L/22-04-ПИИШ.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Хабарова И.А., Валиев Д.С., Чугунов В.А., & Хабаров Д.А.* Современная цифровая фотограмметрия // Международный журнал прикладных наук и технологий «Integral». – 2019. – 4(2). – С. 41-47.
2. *Краснопевцев Б.В.* Фотограмметрия. – М.: УПП "Репрография" МИИГАиК, 2008. – 160 с.
3. *Алтухов, Виктор Григорьевич.* Исследование точности фотограмметрии как метода определения объема объекта // Автоматика и программная инженерия. – 2020. – 2 (32). – С. 69-74.
4. *Безменов В.М.* Применение методов фотограмметрии в вопросах точности обеспечения пространственного и углового положения снимков дистанционного зондирования Земли // Современные проблемы дистанционного зондирования Земли из космоса. – 2020. – 17 (5). – С. 45-52.
5. *Михайлова М.В., Ахмедов А.Н., Шагибалов Р.Р.* Фотограмметрия. Основные принципы и практическое применение // Приборостроение, метрология и информационно-измерительные приборы, и системы. Серия «Естественные и технические науки». – 2018. – № 5.
6. *Грушин С.П., Сосновский И.А.* Фотограмметрия в археологии - методика и перспективы // Теория и практика археологических исследований. – 2018. – № 1 (21).
7. *Катермин В.С.:* Фотограмметрия: 3D-модель из фотографий // Актуальные научные исследования в современном мире. – 2021. – 12-11 (80). – С. 89-94.
8. *уканов А.Н., Цой Е.В., Яковенко А.А., Малий Д.В., Гончаров С.С.* Фотограмметрия в фиксации и анализе локализованной деформации 3d образцов // Современные проблемы и направления развития металловедения и термической обработки металлов и сплавов. – 2023. – С. 168-173.
9. *Ануфриев Владимир Николаевич.* Создание трехмерной модели объектов методами фотограмметрии: аннотация к дипломной работе. – 2022.
10. *Сергеев Н.Е., Самойлов А.Н., Половко И.Ю.* Онтологическое представление фотограмметрических методов для решения задач определения геометрических параметров объектов по предварительно обработанным цифровым изображениям // Вестник АГУ. – 2020. – 3 (266). – С. 34-39.
11. *Бенджо И., Курвилль А., Венсан П.* Глубокое обучение: пер. с англ. А.А. Слинкина. – М.: ДМК Пресс, 2018. – 482 с.
12. *Роджерс А., Хасты М.* Обработка естественного языка с помощью TensorFlow: создание приложений для машинного обучения: пер. с англ. А.В. Логунова. – СПб.: Питер, 2019. – 352 с.
13. *Васильев И.Г., Ларионов С.А.* Модели глубокого обучения для обработки естественного языка: учеб. пособие. – Москва: РУДН, 2020. – 119 с.
14. *Devlin J., Chang M.-W., Lee K., Toutanova K.* BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding // Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. – 2019. – P. 4171-4186.
15. *Liu Y., Ott M., Goyal N., Du J., Joshi M., Chen D., Levy O., Lewis M., Zettlemoyer L., Stoyanov V.* RoBERTa: A Robustly Optimized BERT Pretraining Approach // arXiv preprint arXiv:1907.11692. – 2019.
16. *Radford A., Wu J., Child R., Luan D., Amodei D., Sutskever I.* Language Models are Unsupervised Multitask Learners // OpenAI Blog. – 2019. – Vol. 1, No. 8.
17. *Reimers N., Gurevych I.* Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks // Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). – 2019. – P. 3980-3990.
18. *Conneau A., Kiela D., Schwenk H., Barrault L., Bordes A.* Supervised Learning of Universal Sentence Representations from Natural Language Inference Data // Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. – 2017. – P. 670-680.
19. *Pennington J., Socher R., Manning C.D.* GloVe: Global Vectors for Word Representation / J. Pennington // Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). – 2014. – P. 1532-1543.
20. *Mikolov T., Chen K., Corrado G., Dean J.* Efficient Estimation of Word Representations in Vector Space // arXiv preprint arXiv:1301.3781. – 2013.

REFERENCES

1. *Khabarova I.A., Valiev D.S., Chugunov V.A., Khabarov D.A.* Sovremennaya tsifrovaya fotogrammetriya [Modern digital photogrammetry], *Mezhdunarodnyy zhurnal prikladnykh nauk i tekhnologii «Integral»* [International Journal of Applied Science and Technology "Integral"], 2019, No. 4 (2), pp. 41-47.
2. *Krasnopevtsev B.V.* Fotogrammetriya [Photogrammetry]. Moscow: UPP "Reprografiya" MIIGAiK, 2008, 160 p.

3. Altukhov V.G. Issledovaniye tochnosti fotogrammetrii kak metoda opredeleniya ob"ema ob'yekta [Study of the accuracy of photogrammetry as a method for determining the volume of an object], *Avtomatika i programmnaya inzheneriya* [Automation and Software Engineering], 2020, No. 2 (32), pp. 69-74.
4. Bezmenov V.M. Primeneniye metodov fotogrammetrii v voprosakh tochnosti obespecheniya prostranstvennogo i uglovogo polozheniya snimkov distantsionnogo zondirovaniya Zemli [Application of photogrammetry methods in matters of accuracy of ensuring the spatial and angular position of Earth remote sensing images], *Sovremennyye problemy distantsionnogo zondirovaniya Zemli iz kosmosa* [Sovremennyye problemy distantsionnogo zondirovaniya Zemli iz kosmosa], 2020, No. 17(5), pp. 45-52.
5. Mikhailova M.V., Akhmedov A.N., Shagibalov R.R. Fotogrammetriya. Osnovnye printsipy i prakticheskoye primeneniye. Priborostroyeniye, metrologiya i informatsionno-izmeritel'nye pribory, i sistemy [Photogrammetry. basic principles and practical application. Instrumentation, metrology and information-measuring devices, and systems], *Yestestvennyye i tekhnicheskiye nauki* [Natural and Technical Sciences], 2018, No. 5.
6. Grushin S.P., Sosnovsky I.A. Fotogrammetriya v arkheologii - metodika i perspektivy [Photogrammetry in archeology - methods and prospects], *Teoriya i praktika arkheologicheskikh issledovaniy* [Theory and practice of archaeological research], 2018, No. 1 (21).
7. Katermin V.S. Fotogrammetriya: 3D-model' iz fotografiy [Photogrammetry: 3D model from photographs], *Aktual'nyye nauchnyye issledovaniya v sovremennom mire* [Current scientific research in the modern world], 2021, No. 12-11(80), pp. 89-94.
8. Chukanov A.N., Tsoi E.V., Yakovenko A.A., Maliy D.V., Goncharov S.S. Fotogrammetriya v fiksatsii i analize lokalizovannoy deformatsii 3d obraztsov [Photogrammetry in fixation and analysis of localized deformation of 3d samples], *Sovremennyye problemy i napravleniya razvitiya metallovedeniya i termicheskoy obrabotki metallov i splavov* [Modern problems and directions in the development of metal science and heat treatment metals and alloys], 2023, pp. 168-173.
9. Anufriev V.N. Sozdaniye trekhmernoy modeli ob'yektov metodami fotogrammetrii [Creating a three-dimensional model of objects using photogrammetry methods: abstract to the thesis], 2022.
10. Sergeev N.E., Samoilov A.N., Polovko I.Yu. Ontologicheskoe predstavlenie fotogrammetricheskikh metodov dlya resheniya zadach opredeleniya geometricheskikh parametrov ob'yektov po predvaritel'no obrabotannym tsifrovym izobrazheniyam [Ontological representation of photogrammetric methods for solving problems of determining the geometric parameters of objects from pre-processed digital images], *Vestnik AGU* [Bulletin of ASU], 2020, No. 3(266), pp. 34-39.
11. Bengio Y., Courville A., Vincent P. Glubokoye obucheniye [Deep learning transl. from Engl. A.A. Slinkina.]. Moscow: DMK Press, 2018, 482 p.
12. Rogers A., Hastie M. Obrabotka yestestvennogo yazyka s pomoshch'yu TensorFlow: sozdaniye prilozheniy dlya mashinnogo obucheniya [Natural language processing with TensorFlow: creating applications for machine learning lane: transl. from Engl. Logunova A.V.]. St. Petersburg: Peter, 2019, 352 p.
13. Vasiliev I.G., Larionov S.A. Modeli glubokogo obucheniya dlya obrabotki yestestvennogo yazyka: ucheb. posobiye [Deep learning models for natural language processing: textbook]. Moscow: RUDN, 2020, 119 p.
14. Devlin J., Chang M.-W., Lee K., Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2019, pp. 4171-4186.
15. Liu Y., Ott M., Goyal N., Du J., Joshi M., Chen D., Levy O., Lewis M., Zettlemoyer L., Stoyanov V. RoBERTa: A Robustly Optimized BERT Pretraining Approach, *arXiv preprint arXiv:1907.11692*, 2019.
16. Radford A., Wu J., Child R., Luan D., Amodei D., Sutskever I. Language Models are Unsupervised Multitask Learners, *OpenAI Blog*, 2019, Vol. 1, No. 8.
17. Reimers N., Gurevych I. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks, *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 2019, pp. 3980-3990.
18. Conneau A., Kiela D., Schwenk H., Barrault L., Bordes A. Supervised Learning of Universal Sentence Representations from Natural Language Inference Data, *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, 2017, pp. 670-680.
19. Pennington J., Socher R., Manning C.D. GloVe: Global Vectors for Word Representation, *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2014, pp. 1532-1543.
20. Mikolov T., Chen K., Corrado G., Dean J. Efficient Estimation of Word Representations in Vector Space, *arXiv preprint arXiv:1301.3781*, 2013.

Статью рекомендовал к опубликованию д.т.н., профессор В.В. Курейчик.

Козловский Александр Вячеславович – Южный федеральный университет; e-mail: kozlovskiy@sfedu.ru; г. Таганрог, Россия; тел.: +79888980779; кафедра вычислительной техники; ассистент.

Мельник Эдуард Всеволодович – e-mail: evmelnik@sfedu.ru; тел.: +78632184000, кафедра вычислительной техники; профессор.

Самойлов Алексей Николаевич – e-mail: asamoylov@sfedu.ru; тел.: +78632184000; кафедра вычислительной техники; зав. кафедрой.

Kozlovskiy Alexander Vyacheslavovich – Southern Federal University; e-mail: kozlovskiy@sfedu.ru; Taganrog, Russia; phone: +79888980779; the Department of Computer Engineering; assistant.

Melnik Eduard Vsevolodovich – e-mail: evmelnik@sfedu.ru; phone: +78632184000; the Department of Computer Science; professor.

Samoylov Aleksey Nikolayevich – e-mail: asamoylov@sfedu.ru; phone: +78632184000; the Department of Computer Science; head of the Department.

УДК 621.317.3

DOI 10.18522/2311-3103-2024-3-252-265

А.А. Яковлев, М.Ю. Серов, Р.В. Сахабудинов, А.С. Голосий**МОДЕЛИРОВАНИЕ ЭЛЕКТРИЧЕСКИХ ХАРАКТЕРИСТИК РАЗРЯДНОГО КОНТУРА ВЫСОКОВОЛЬТНОГО СТЕНДА МОЛНИЕВОГО РАЗРЯДА**

Устойчивое прохождение ракетой-носителем атмосферного слоя, где могут возникать разряды электричества, обеспечивается не только организационными мерами, но и предварительными наземными испытаниями. Государства, осуществляющие вывод космических аппаратов на орбиту, располагают специальным стендовым оборудованием. Сложилась определенная система взглядов, реализованная в стандартах и других технических документах, требования которых стали обязательны к выполнению. Настоящая работа является логическим продолжением исследований, связанных с созданием высоковольтного стенда молниевого разряда (ВСМР), разрабатываемого для испытаний изделий ракетно-космической техники. Основной задачей ВСМР является генерация заданных электрических (или электромагнитных) импульсов, имитирующих воздействие молниевого разряда на элементы конструкции ракето-носителя. В высоковольтной части ВСМР состоит из четырех генераторов импульсных токов (ГИТ-А, ГИТ-В, ГИТ-С и ГИТ-Д), последовательно подключаемых к нагрузке для создания определённой формы общего токового импульса. Типовые нагрузки включают: стол заземления высоковольтный, стойку вертикальную, приспособление на испытание пробоем. Задачей настоящего этапа работ явилась проверка параметров токового импульса, возникающего при разряде генератора импульсных токов ГИТ-А на калибровочную нагрузку, в качестве которой принимается стол заземления высоковольтный. В статье представлены результаты расчетов параметров импульса компоненты «А» в разрядном контуре в ходе развития процесса генерации импульса: до момента закоротки емкостного накопителя и от момента закоротки. Разрядное устройство «кроубар» позволяет подключать нагрузку по двухконтурной схеме в момент максимума разрядного тока. Разработаны аналитические зависимости обеих эквивалентных электрических схем подключения контуров. Дифференциальные уравнения решены численным методом, получены графики изменения тока и напряжения колебательного импульса «А» в незакороченном и закороченном контурах. Моделирование позволило оценить динамические характеристики исследуемого контура при его функционировании в одном из самых быстропотекающих и энергоемких режимов работы. В целом коммутация разрядного контура на стенд заземления высоковольтный с подобранными параметрами подтверждает работоспособность ВСМР и достижение удовлетворительных характеристик заданного токового импульса, реализуемого ГИТ-А.

Молниевый разряд; высоковольтный стенд молниевого разряда; генератор импульсных токов; закоротка; емкостной накопитель; импульс; кроубар; электрический контур; максимум тока; индуктивность.