

14. Hartmann W., Havlena M., Schindler K. Recent developments in large-scale tie-point matching, *ISPRS Journal of Photogrammetry and Remote Sensing*, 2016, Vol. 115, pp. 47-62.
15. Zhao Q., Feng W., Wan L., Zhang J. Sphorb: A fast and robust binary feature on the sphere, *International journal of computer vision*, 2015, Vol. 113, pp. 143-159.
16. Torii A., Havlena M., Pajdla T. From google street view to 3d city models, in *2009 IEEE 12th international conference on computer vision workshops, ICCV Workshops*, 2009, pp. 2188-2195.
17. Tian Y., Fan B., Wu F. L2-net: Deep learning of discriminative patch descriptor in euclidean space, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 661-669.
18. Murrugarra-Llerena J., da Silveira T.L., Jung C.R. Pose estimation for two-view panoramas based on keypoint matching: A comparative study and critical analysis, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 5202-5211.
19. Jiang S., Jiang C., Jiang W. Efficient structure from motion for large-scale UAV images: A review and a comparison of sfm tools, *ISPRS Journal of Photogrammetry and Remote Sensing*, 2020, Vol. 167, pp. 230-251.
20. Mouragnon E., Lhuillier M., Dhome M., Dekeyser F., Sayd P. Real time localization and 3d reconstruction, in *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2006, Vol. 1, pp. 363-370.
21. Scaramuzza D., Martinelli A., Siegwart R. A flexible technique for accurate omnidirectional camera calibration and structure from motion, in *Fourth IEEE International Conference on Computer Vision Systems*, 2006, pp. 45-45.

Статью рекомендовал к опубликованию д.т.н., профессор А.В. Боженюк.

**Морев Кирилл Иванович** – Акционерное общество Научно-конструкторское бюро вычислительных систем (АО НКБ ВС); e-mail: morev-ki@ya.ru; г. Таганрог, Россия; ; программист.

**Ледерер Петр Алексеевич** – e-mail: plederer25@gmail.com; программист.

**Morev Kirill Ivanovich** – Joint Stock Company "Scientific Design Bureau of Computing Systems" (JSC SDB CS); e-mail: morev-ki@ya.ru; Taganrog, Russia; programmer.

**Lederer Pyotr Alekseevich** – e-mail: plederer25@gmail.com; programmer.

УДК 28.23.24: 28.23.27: 004.82

DOI 10.18522/2311-3103-2024-1-285-299

**С.М. Соколов**

## **АНАЛИЗ СЦЕНЫ В ИНФОРМАЦИОННЫХ СИСТЕМАХ ПОДВИЖНЫХ РОБОТОТЕХНИЧЕСКИХ КОМПЛЕКСОВ**

*Современные роботы способны выполнять все более сложные задачи, которые обычно требуют высокой степени взаимодействия с окружающей средой, в которой им приходится работать. Как следствие, роботизированные системы должны обладать глубокими и конкретными знаниями о своих рабочих пространствах, которые выходят далеко за рамки простого представления показателей, которое роботизированная система может создать с помощью приёмов обработки только зрительных данных, например в задаче одновременной локализации и картографирования (SLAM). Анализ сцены является связующим звеном между распознаванием объектов и знанием об окружающем мире и в том или ином виде присутствует в процессе извлечения из зрительных данных информации, необходимой для решения конкретной задачи. В статье представляется систематизированный подход к обеспечению анализа сцены бортовыми СТЗ. Рассматриваются технологии анализа сцены как составной части повышения степени автономности подвижных РТК. Ряд технологий только предстоит освоить и воплотить в жизнь, но общая структура позволяет постепенно углублять анализ сцены на борту РТК, тем самым повышая степень автономности без коренной переделки бортовой информационно-управляющей системы и СТЗ, как ключевой части информационного обеспечения. Информация, извле-*

чѐнная из зрительных данных, интегрируется в многослойную карту, обеспечивая высокоуровневое представление окружающей среды, которое воплощает в себе знания, необходимые робототехническому комплексу для реального выполнения сложных задач. Многослойная карта представляет собой форму хранения знаний об окружающей обстановке и объектов в ней. Эта карта объединяет пространственную иерархию объектов и мест с семантической иерархией понятий и отношений. Описываются структуры для представления данных в различных слоях этой карты и механизмы их использования. В частности, для описания маршрутов движения РТК, используются принципы интерпретирующей навигации, для представления информации об условиях функционирования и объектах интереса структуры сигнатур. В основе программной реализации предлагаемых механизмов используется унифицированный подход на основе программного каркаса СТЗ реального времени. Приводятся примеры использования описанных технологий при решении задач информационного обеспечения целенаправленных перемещений наземных РТК.

Степень автономности; система технического зрения; анализ сцены; многоуровневые когнитивные карты; конфигурационное пространство; каркас программного обеспечения СТЗ реального времени; онтологии робототехники; сигнатура.

S.M. Sokolov

### SCENE ANALYSIS IN MOBILE INFORMATION SYSTEMS ROBOTIC COMPLEXES

Modern robots are capable of performing increasingly complex tasks that usually require a high degree of interaction with the environment in which they have to work. As a result, robotic systems must have deep and specific knowledge about their workspaces, which go far beyond the simple representation of indicators that a robotic system can create using visual data processing techniques, for example, in the task of simultaneous localization and mapping (SLAM). Scene analysis is the link between object recognition and knowledge about the world around us and is present in one form or another in the process of extracting information from visual data necessary to solve a specific task. The article presents a systematic approach to providing on-board STZ analysis of the scene. The technologies of scene analysis are considered as an integral part of increasing the degree of autonomy of mobile RTCs. A number of technologies have yet to be mastered and implemented, but the overall structure allows you to gradually deepen the analysis of the scene on board the RTK, thereby increasing the degree of autonomy without radically redesigning the on-board information management system and STZ, as a key part of information support. The information extracted from the visual data is integrated into a multi-layered map, providing a high-level representation of the environment, which embodies the knowledge necessary for a robotic complex to actually perform complex tasks. A multi-layered map is a form of storing knowledge about the environment and the objects in it. This map combines a spatial hierarchy of objects and places with a semantic hierarchy of concepts and relationships. The structures for representing data in various layers of this map and the mechanisms for their use are described. In particular, to describe the routes of the RTK, the principles of interpretive navigation are used to provide information about the operating conditions and objects of interest of the signature structure. The software implementation of the proposed mechanisms is based on a unified approach based on the real-time STZ software framework. Examples of the use of the described technologies in solving the problems of information support for targeted movements of ground RTCs are given..

Degree of autonomy; vision system; scene analysis; multilevel cognitive maps; configuration space; real time vision system framework; robotics ontologies; signature.

**Введение.** Учет сложности среды, окружающей робота, необходим для эффективного применения роботов с повышенной степенью автономности. Однако по-прежнему существует разрыв в ожиданиях пользователей и функциональности подобных роботов. Одним из ограничивающих факторов современных технологий является плохое восприятие окружающей среды, что резко влияет на то, что робот может понять о своем окружении. С одной стороны, это привело к значительным усилиям по совершенствованию сенсорных устройств и возможностей восприятия. С другой стороны, проблема решается путем оказания роботу поддержки в приобретении знаний и осведомленности о рабочей среде. В настоящее время при суперви-

зорном управлении рассматривается подход к анализу сцены, основанный на построении высокоуровневого представления окружающей среды, которое включает как геометрическую, так и символическую информацию. Выполняется разделение обязанностей. Робот отвечает за построение метрической карты окружающей среды, в то время как человек-оператор предоставляет пространственные подсказки о каждом возможном интересующем объекте, который должен быть включен в карту. Это направление можно назвать методом SLAM (одновременной локализации и картографирования) для построения карты окружающей среды, дополненной способностью интегрировать семантическую информацию, предоставленную пользователем.

Планирование задач для мобильных роботов обычно основывается исключительно на пространственной информации и поверхностных знаниях предметной области, таких как метки, прикрепленные к объектам и местам. Хотя пространственная информация необходима для выполнения основных операций робота (навигации и локализации), использование более глубоких знаний предметной области имеет решающее значение для наделяния робота более высокой степенью автономии и интеллекта, что особенно важно при перемещениях в естественных средах.

В настоящей работе подчеркивается тесная связь анализа сцены и степени автономности РТК. На основе анализа имеющихся публикаций и собственного опыта по исследованиям и разработке РТК с ПСА делается попытка систематизации технологий, которые могут быть применены при анализе сцены в бортовой информационно-управляющей системе робота с повышенной степенью автономности. Формируется общая схема анализа сцены, указывающая на разделение функций между оператором и РТК и открытая для наполнения и редакции по мере повышения степени автономности. Приводятся примеры использования этой схемы для решения задач информационного обеспечения целенаправленных перемещений наземных РТК с ПСА.

**Известные решения в области анализа сцены.** Анализ сцен имеет долгую историю, практически совпадающую с историей компьютерного зрения [1–3]. Если первые исследования в этой области рассматривали модельные сцены и акцентировали логический уровень рассмотрения [4], то в настоящее время на первый план выходит анализ естественных сцен с целью извлечения информации, необходимой, в частности, для систем управления автономных робототехнических комплексов [5, 6]. Хорошая связь степени автономности РТК и необходимости представления знаний об окружающей среде отражена в так называемом «пространстве автономии» (рис. 1).



Рис. 1. Пример показателей для определения уровня автономности, предложенные группой ALFUS («пространство автономии»)

Эта схема показывает связь желаемой/проектируемой степени автономности со сложностью и объёмом анализа сцены и необходимость распределения работ по анализу сцены между априорной подготовкой инфраструктуры области функцио-

нирования, оператором и бортовой информационно-управляющей системой. Подробно эти вопросы разбираются в [7]. Здесь отметим только тот факт, что автономность беспилотных средств касается множества технических областей. Сложность задачи и адаптивность к окружающей среде являются одними из ключевых аспектов. Для достижения целей, ориентированных на конкретные задачи, необходимо, чтобы любая значимая сущность окружающей среды (т.е. объекты, области и т.д.) могла обрабатываться как семантическое знание, чтобы получить то, что обычно называют семантической картой. Согласно Nuchter и Hertzberg [8]: семантическая карта для мобильного робота – это карта, которая содержит, в дополнение к пространственной информации об окружающей среде, присвоения отображаемых объектов объектам известных классов.

До последнего времени, в публикациях анализ сцены рассматривался в контексте исследований компьютерного зрения и интеллектуальных технологий. В робототехнических задачах анализ сцены был поглощён методами одновременного определения своего положения и построения карты (SLAM). Тем самым, делая акцент на получении и работе с количественным, метрическим описанием окружения. Сейчас положение меняется и намечается активный переход к развитию методов SLAM, объединяющих метрические и семантические карты области функционирования, что отводит анализу сцены полноценную роль, служащую повышению степени автономности. Исследования по совместному использованию зрительных данных и знаний об объектах внешнего мира, представленных, в частности, в текстовой форме, проводились ещё до «бума» нейросетевых трансформеров. В этом направлении активно велись работы по дополнению традиционных навигаторов многоиерархическим представлением городских карт. Передовые информационные системы для путешественников (ATISs) – это технология, которая имеет дело с интеллектуальными транспортными приложениями, направленными на предоставление пользователям жизненно важной информации о дорожном движении [9]. Исходно, такие приложения были ориентированы на интерактивный режим работы с пользователем, но заложенные механизмы позволяют двигаться по направлению повышения степени автономности транспортного средства и переходить к супервизорному управлению, используя дополнительную к зрительной информации в бортовых системах принятия решения. Например, по сочетанию нескольких внешних признаков и априорной информации об области функционирования, принимать решение о выборе траектории движения с учётом не только геометрической, но и несущей проходимости.

Ключевым моментом является то, как правильно представить и объединить семантические знания и геометрическую информацию. Связанную с этим работу можно разделить на две основные категории, отличающие полностью автоматические методы, не предполагающие взаимодействия с человеком, от полуавтоматических подходов, когда пользователь участвует в процессе получения знаний, поддерживая работу в построении семантической карты. В настоящее время при супервизорном (полуавтоматическом) управлении рассматривается подход к анализу сцены, основанный на построении высокоуровневого представления окружающей среды, которое включает как геометрическую, так и символическую информацию. Выполняется разделение обязанностей. Робот отвечает за построение метрической карты окружающей среды, в то время как человек-оператор предоставляет пространственные подсказки о каждом возможном интересующем объекте, который должен быть включен в карту [10].

Полностью автоматические методы можно далее разделить на три группы методов. Первая группа направлена на извлечение особенностей окружающей среды из лазерных карт для поддержки маркировки. Эти подходы включают определение атрибутов исследуемых пространств [11] и использование дополнения топологических карт семантическими знаниями с использованием привязки [12]. Второй набор методов использует классификацию и кластеризацию для автоматиче-

ской сегментации и маркировки метрических карт. Например, в работе Nuchter et al. [8], знания об окружающей среде извлекаются путем маркировки трехмерных точек с помощью разности градиентов между соседними точками. Генерация двумерных топологических карт из метрических карт была описана в [13, 14] (с использованием AdaBoost), в Brunskill et al. [15] (с использованием спектральной кластеризации) и в [16] (с использованием случайных полей Вороного). Третий набор методов распознавания объектов и категоризации мест основан на извлечении визуальных признаков, таких как в [17], или комбинации данных, полученных камерой RGB-D, таких как в [18]. Несмотря на значительный прогресс, достигнутый в полностью автоматизированном семантическом отображении, подход по-прежнему страдает от ошибок и отсутствия общности.

В полуавтоматических подходах оператор играет роль супервизора или оператора, обычно взаимодействующего с помощью естественного языка. В Diosi et al. [17] интерактивная процедура SLAM и сегментация водораздела используются для создания контекстной топологической карты. Зендер и др. [18] описали систему, способную создавать концептуальные представления о внутренней среде, созданной человеком. Роботизированная платформа обладает априорными знаниями о пространственных концепциях, используемыми для создания внутреннего представления окружающей среды, получаемого с помощью датчиков низкого уровня. Роль оператора на протяжении всего процесса сбора данных заключается в поддержке робота в нанесении меток на места. Пронобис и Йенсфельт [19] представляют многоуровневый алгоритм семантического отображения, объединяющий информацию о существовании объектов и семантических свойствах пространства. Вводимые пользователем данные, если они предоставлены, интегрируются в систему в качестве дополнительных свойств существующих объектов. Наконец, Нието-Гранда и др. [20] используют дополненное человеком отображение, основанное на многомерной вероятностной модели, для привязки пространственной области к семантической метке. Оператор поддерживает робота в этом процессе, давая ему инструкции по выбору меток. Немногие подходы нацелены на более продвинутую форму совместной работы, когда оператор активно сотрудничает с роботом для создания семантической карты не только для категоризации мест и маркировки, но и для распознавания объектов и позиционирования. Из-за сложности такого взаимодействия, чтобы оно не приводило к утомительным усилиям для неопытного оператора, предлагаются мультимодальные взаимодействия, позволяющее естественным образом работать с различными типами информации. Например, Kruijff et al. [21] внедряют систему для улучшения процесса картографирования с использованием естественного языка; Ранделли и др. [22] предлагают богатое мультимодальное взаимодействие, включающее речь, жесты и видение для семантической маркировки ориентиров окружающей среды.

Несмотря на тот факт, что взаимодействие с окружающей средой имеет первостепенное значение для большинства приложений робототехники современные системы, реализующие подходы SLAM не способны обеспечить тесное понимание на высоком уровне геометрии и семантики окружающего мира; разработка таких представлений должна быть ориентирована на задачи и в настоящее время это гибкая структура, позволяющая связать задачу с оптимальными представлениями отсутствием. Разработку такой структуры может принести только совместная работа сообщества робототехников, искусственного интеллекта и компьютерного зрения. По мнению авторов обзора проблем SLAM [23] машинное обучение на основе глубоких нейросетей будет иметь влияние на SLAM, но при этом следует учитывать особенности и ограничения этого направления. В этих технологиях эффективна классификация хорошо известных объектов, достигая уровня производительности около 80% и более [24]. При переходе в системах SLAM к более семантически значимым картам эти инструменты «черного ящика» могут ускорить этот переход. Но успехи в глубоком обучении в основном связаны с длительным вре-

менем обучения на суперкомпьютерах и выводе на специализированные графические процессоры для одноразового результата. Задача для исследователей SLAM (или даже для всех, кто хочет внедрить впечатляющие результаты в своей системе) заключается в том, как обеспечить достаточную вычислительную мощность во встроженных системах. Необходимо исследовать применимость этих результатов к меньшим масштабам вычислительных средств. Еще более серьезной и важной проблемой является онлайн-обучение. SLAM системы обычно работают в открытом мире с непрерывным наблюдением, где могут встретиться новые объекты и сцены. Но на сегодняшний день глубокие сети обычно обучаются на некоторых фиксированных сценариях и множестве данных (dataset), скажем, с фиксированным количеством классов объектов. Существующие сети, как правило, обучаются на огромном массиве размеченных данных, однако не всегда можно гарантировать, что подходящий набор данных существует или его практически можно использовать. Серьезной задачей является использование мощности глубоких сетей в сценарии с однократным или нулевым предъявлением (т. е. с одним или даже нулевым обучающим примером нового класса). Предварительная информация о сцене обеспечивает значительное повышение эффективности систем SLAM. Примеры в литературе на сегодняшний день включают известные объекты или предварительные знания об ожидаемой структуре в сцене вроде плавности как в [25], Манхэттен ограничениях, как в [26.], или даже ожидаемые отношения между объектами [27]. Очевидно, что глубокое обучение способно выделять такие предварительные знания для конкретных задач, таких как оценка меток сцены или глубины сцены. Как лучше всего извлечь и использовать эту информацию является серьезной открытой проблемой. Методы SLAM хорошо разбираются в математике геометрии сцены и встает вопрос - как соединить это хорошо понятную геометрию с выходами глубокой сети? Особая задача, которую необходимо решить, состоит в том, чтобы охарактеризовать неопределенность оценок, полученных на основе глубокой сети. Исследователи SLAM предлагают сложный контекст для изучения потенциала связи между архитектурами глубокого обучения и рекурсивными оценка состояния в крупномасштабных графических моделях, например, Кришан и другие [28] недавно предложили ряд Калмановских фильтров. Возможно, однажды станет возможным создать сквозная SLAM-система, использующая глубокую архитектуру, без явного моделирования функций, ассоциация данных и т. д.

В последние годы активное развитие получили методы объединения обработки зрительных данных совместно с текстами, описывающими изображения, с помощью генеративных нейросетей. Так, в беспилотных пассажирских перевозках голосовые запросы обрабатываются не напрямую алгоритмом беспилотного вождения, а опосредовано. Запрос при помощи технологии анализа голоса переводится в некоторый запрос – обычно признаковое представление запроса. После этого, уже по этому запросу, например, выбирается адрес и до него строится маршрут при помощи карты. Из отдельных модулей - эмбедингов формируется обобщенное представление, решающие конкретную подзадачу. Каждый модуль представляет собой некоторую сеть-трансформер. Другой пример такого объединения приводится в работе [29], где на основе общего представления осуществляется поиск действий в видео по описанию и, наоборот, – составление описания по видео. Пример блок-схемы алгоритма детектирования объектов на изображениях по их описаниям представлен на рис. 2.

Замечания, высказанные в обзоре по методам SLAM и опыт нашей группы по исследованию нейросетевых подходов в робототехнике говорит о том, что применение нейросетевых решений требует очень взвешенного рассмотрения.

**Технологии анализа сцены и наш подход.** В секторе робототехники ИПМ им. М.В. Келдыша РАН анализ сцены рассматривается как инструмент познания, который РТК использует в процессе функционирования в реальном мире. Степень, детальность этого анализа определяется поставленной и решаемой РТК задачей.

Анализ сцены является важной составляющей в алгоритмическом обеспечении СТЗ в составе бортовой информационно-вычислительной системы (БИУС). Общая схема создания этой части алгоритмического обеспечения следующая. После фиксации на очередной итерации в конфигурационном пространстве [30] координат по оси моделей и составу аппаратной поддержки, выполняется исследование интервала возможного алгоритмического обеспечения по оси алгоритмов. Это исследование выполняется с помощью имитационного моделирования, компоновки алгоритмов, реализованных в унифицированном каркасе ПО СТЗ РВ [31]. Выполняется исследование разделения работ между оператором и БИУС с использованием технологий многоуровневых карт и сигнатур.

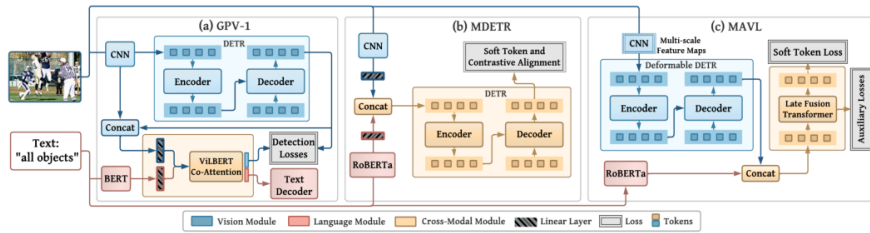


Рис. 2. Блок-схема метода детектирования объектов по их описаниям на основе нейросетей трансформеров

В качестве носителей знаний об окружающем РТК пространстве при разработке алгоритмического обеспечения выступает многослойная карта. Здесь представлены пространственные и функциональные свойства окружающей среды, которые позволяют роботу воспринимать окружающую среду так же, как это делает человек. Такая карта позволяет роботу воспринимать окружающую среду геометрически, а также использовать семантические представления, что может способствовать эффективному планированию и взаимодействию человека и робота. Как показано на рис. 3, многослойная карта состоит из трех слоев: геометрической карты внизу, топологической карты посередине для представления распределенных областей на графике и семантической карты сверху для представления свойств и классов объектов интереса.

<p><b>Семантическая карта:</b> граф смысловых связей объектов интереса</p>
<p><b>Топологическая карта:</b> граф пространственных связей объектов интереса</p>
<p><b>Физическая, геометрическая карта:</b> Форматы, слои отображений сенсорных данных и используемой ГИС</p>

Рис. 3. Схема слоёв многослойной карты для представления данных о пространстве функционирования РТК

Семантическое картирование – это постепенный процесс сопоставления релевантной информации о мире (т.е. пространственной информации, временных событиях, агентах и действиях) с формальным описанием, поддерживаемым логическим механизмом, с целью обучения пониманию, сотрудничеству и коммуникации.

В последние годы семантическое картографирование стало очень активной областью исследований. Такой растущий интерес продиктован идеей о том, что если роботы могут понимать среду, в которой живут люди, и то, как они действуют в ней, они также могут сотрудничать и действовать (т.е. обладать более когнитивным поведением). Такое представление является необходимым условием для совместной работы двух или более агентов. Семантические карты являются необходимым условием объединения разнородных и разнородных групп РТК.

Для того, чтобы определить карту как “семантическую”, требуется, чтобы знания были представлены подходящим образом. Необходима формализация минимально общей структуры представления, которая должна быть реализована в семантической карте. Это представление должно играть роль общего интерфейса между всеми семантическими картами и может быть легко расширено или специализировано по мере необходимости. Это представление определяется как тройное  $SM = \langle R, M, P \rangle$ , где:

- ◆  $R$  – глобальная система отсчета, в которой выражены все элементы семантической карты;
- ◆  $M$  – набор геометрических элементов, полученных в качестве необработанных данных датчиков. Они выражены в системе отсчета  $R$  и описывают пространственную информацию в математической форме.  $M_s \subseteq M$  – подмножество семантически значимых элементов;
- ◆  $P$  – это набор предикатов, среди которых  $-a(X, Y)$  и  $instance-of(X, Y)$  являются обязательными.

В нашей практике на семантическом уровне используем граф связей ОИ (на основе онтологий предметной области). Топологический уровень – графы информационной эквивалентности, районы информационной эквивалентности, в соответствии с концепцией интерпретирующей навигации [32]. Геометрический/физический (количественный/метрический) уровень форматы и слои ГИС «Панорама», результаты построения карты по данным датчиков.

Карты занятости на основе сетки рабочего пространства могут создавать точное пространственное представление для локализации роботов и безопасной навигации. Но сложность планирования на основе карты сетки часто очень велика в крупномасштабных средах. Трун [33] представил метрико-топологический картографический подход для навигации мобильных роботов внутри помещений, который может использоваться гораздо эффективнее, но также согласуется с картой на основе сетки. *О. М. Mozos*, и др. [34?] в разработали эффективную систему навигации роботов с гибридной метрико-топологической картой. Их система использует алгоритм Дейкстры для планирования на топологической карте с позиционным графом, и их результат продемонстрировал, что планирование на основе гибридной карты более эффективно, чем планирование на основе метрик. Топологические карты разделяют регионы поверх карты-сетки. Мозос и др. [34] представили метод извлечения топологической семантической карты из метрической с использованием контролируемого обучения. Граф может быть определен как  $G = \{N, E, S, P\}$ , где:

- ◆  $N$  – набор узлов, таких как районы информационной эквивалентности (имеющие своё информационное описание в терминах ИН). Каждый узел имеет уникальное имя.
- ◆  $E$  – это набор ребер, который представляет ограничение (связи) – переходы между узлами.
- ◆  $S$  относится к набору текущей области робота с учетом данных датчиков и одометрии на геометрической карте (текущая, определяемая по датчикам позиция).
- ◆  $P$  – это набор глобальных положений каждого узла.



С помощью топологической карты робот или оператор может выбрать различные навигационные модули в разных регионах – например, в помещении или в хорошо организованном городском пространстве навигация управляется планировщиком на основе карты-сетки; как только осуществится выход в естественную среду, модуль навигации в естественной среде возьмет управление на себя. Верхний уровень – это семантическая карта, которая включает в себя пространственную концепцию сред, основанную на онтологии, связанную с топологической картой и геометрической картой. Это дает единообразное описание сред с метаданными, а также контекст объектов внутри сред. Это обеспечивает связь между пространственной информацией и концепцией человека (семантикой), которая позволяет расположить и естественная коммуникация человека и робота. Основываясь на работе [36], база знаний онтологии среды объединяется с информацией о топологии и геометрии. Онтология содержит таксономии и взаимосвязи свойств сред, которые позволяют делать выводы и рассуждения о гетерогенных средах.

Перспективным направлением в создании и описании пространственных представлений является применение четко определенных концептуальных онтологий. Применение этого строгого подхода упрощает процесс преобразования постановок задач на естественном языке во внутреннее представление робота и наоборот. В контексте навигационных роботов знание предметной области чаще всего требуется для моделирования представления пространства, движения и действия. Литература по когнитивной робототехнике в настоящее время становится все более насыщенной формально смоделированными представлениями пространственных знаний. Базовые расчеты регионального представления недавно были дополнены более целенаправленными моделями информации о типе маршрута [37] и пространственных отношениях. Хотя такие формальные представления в настоящее время неоднородны по своей природе, существует стремление к унифицированному подходу с использованием верхних пространственных онтологий, ниже которых могут быть разработаны более конкретные онтологии предметной области для отдельных мобильных роботов. Такое четко определенное представление пространственных концепций должно затем позволить реализовать более конкретные и расширяемые системы, которые связывают язык с пространственными знаниями.

На перспективу активного использования РТК с ПСА западные коллеги разрабатывают онтологии верхнего уровня робототехники и онтологии прикладных областей [38]. В России также начались работы в этом направлении, готовится первый стандарт по онтологиям в робототехнике. Мы же в своих экспериментах, не имея достаточной базы данных для формирования онтологии прикладной области, используем смешанный подход: отталкиваясь от функциональных возможностей РТК для описания сценариев его работы используем термины из известного нам тезауруса предметной области (лексикона конечных пользователей). У каждого объекта интереса (ОИ) есть свойства (описанные в сигнатуре ОИ). Онтологическое описание среды реализовано с помощью семантических формализмов OWL. С помощью описания OWL, например, мы можем определить семантическое представление области функционирования РТК. В основе, мы определяем следующие понятия для представления окружающей среды. Класс типа района информационной эквивалентности (РИЭ) относится к разделению областей функционирования по типам ориентиров и других объектов интереса, а также типам опорной проходимости. У каждого объекта интереса есть класс, такие свойства, как ориентация, подкласс, пространственные отношения с другими объектами. Пространство разделено на районы инф эквивалентности со своими специфическими способами выделения ориентиров области. При наличии семантики, созданной OWL, мы можем использовать структуру рассуждения о семантической

карте, которая включает интерфейс Prolog для запроса базы знаний онтологии OWL в ROS. Помимо таксономий свойства среды, мы также можем определить онтологию задачи с помощью OWL, т.е. размещение чего-либо в заданном районе приведет к последовательности рассуждений и действий.

Анализ сцены состоит из нескольких подзадач. Первой из этих подзадач является извлечение информации из зрительных данных. Направляется это извлечение сигнатурами объектов интереса. Одним из основных алгоритмов, используемых на старте сегментации выступает алгоритм сегментация изображения сцены на однородные области. Совместно с априорной информацией о возможных объектах интереса в поле зрения результаты сегментации позволяют более точно и надёжно исследовать объекты интереса, сосредотачивая внимание на тех областях, в которых могут содержаться дополнительные характеристики этих объектов (границы, форма). В качестве наиболее эффективных в составе БИУС были выбраны графовые алгоритмы. Программное обеспечение для работы с графовыми структурами унифицировано в рамках каркаса программного обеспечения.

В дополнение к описанной структуре многоуровневых карт, мы используем специализированные сигнатуры. Эти структуры представляют собой хранилища, позволяющие объединить информацию об объектах интереса с информацией в слоях семантической карты на борту (проверка релевантности реальности модели или построение этой модели) и конструктивно использовать различные методы обработки зрительных данных. Реализовывать целенаправленный сбор зрительных данных, управляемый моделью условий и объектов интереса. В структуре сигнатур, предусматривается помимо перечня необходимых вычислений, возможные действия по выбору поля зрения или необходимости изменения точки наблюдения. Эти данные требуют согласованного использования многоуровневых карт. Механизм оперативной автоматической работы с такими представлениями находится в стадии разработки. В практических задачах реализуется статическая фиксация действий РТК, а для оперативного использования знаний предполагается участие оператора. Например, указание новых ориентиров по ходу движения. По мере развития алгоритмической базы, описанная схема реализации анализа сцены допускает постепенное увеличение степени автономности РТК без коренной перестройки всей архитектуры программного обеспечения бортового вычислительно-управляющего комплекса. Эта схема хорошо продолжает подход, предложенный в монографии [39]. Можно провести параллели в используемых понятиях. Концептуализация предметной области – построение онтологий, типовая ситуация – прецедент и реализующие его сценарии, формирование бортовой оперативно-советующей экспертной системы для типовой ситуации – формирование сигнатур объектов интереса и условий наблюдения в соответствии с решаемой задачей и подготовленными сценариями. В основе программной реализации предлагаемых механизмов используется унифицированный подход на основе программного каркаса СТЗ реального времени [31].

С использованием описанных технологий в нашем секторе был разработан макетный образец программно-аппаратного комплекса для информационного обеспечения целенаправленных перемещений РТК с ПСА. Макетный образец позволил реализовать ряд экспериментов. В частности, было опробовано решение задачи определения траектории движения РТК при движении по смешанной: урбанистической и естественной среде в супервизорном режиме. По маршруту следования подвижного РТК оператор по мере запросов от РТК в полях зрения видеокамер, которыми оснащена бортовая СТЗ, указывал объекты, которые могут быть использованы в качестве ориентиров при решении навигационной задачи (эти данные являются независимым информационным каналом, дополняющим визуальную одометрию, показания инерциальной и спутниковой навигационных

систем). Бортовая информационно-управляющая система проверяет надежность выделения указанных объектов и, в случае положительного результата использует эти объекты в качестве ориентиров, следя за ними и определяя положение РТК относительно них. Когда указанные ориентиры приближаются к выходу из полей зрения, БИУС запрашивает оператора об указании новых ориентиров. Как показали эксперименты, для надежного определения (с точностью до десятка сантиметров) положения РТК, оснащённого одной вперёдсмотрящей стереосистемой и двигающегося со скоростью 10 м/с, без использования других навигационных средств, требуется в среднем не менее 6 ориентиров на 100 метров пути. На 600 метровой замкнутой траектории рассогласование координат, определённых по показаниям высокоточного интегрированного (с БИНС) GPS приёмника и определённых по показаниям визуального одометра, скорректированным данными по определению положения РТК относительно ориентиров, оставалось в пределах 1 м. Кроме этого, проводились эксперименты по детектированию заданных классов препятствий по ходу движения РТК. Препятствия описывались сигнатурами, содержащими, как характеристики образов этих объектов, так и указание на то, какие параметры препятствия следует определить осмотром или переключением на другое поле зрения, чтобы принять решение о преодолении или объезде. В существующей реализации эти действия выполнялись в старт-стопном режиме с привлечением оператора. В целом, описанная схема и выбранные технологии анализа сцены, показали свою успешную работоспособность с возможностью наращивания степени автономности.

**Заключение.** Как показал проведённый анализ доступных публикаций и известного опыта разработок, несмотря на тот факт, что взаимодействие с окружающей средой имеет первостепенное значение для большинства приложений робототехники современные системы, реализующие подходы SLAM не способны обеспечить понимание на высоком уровне геометрии и семантики окружающего мира. Разработка таких представлений должна быть ориентирована на задачи и в настоящее время это гибкая структура, позволяющая связать задачу с оптимальными представлениями отсутствует. Совершенствование технологий анализа сцены, как и повышение степени автономности РТК требуют объединённых усилий как робототехников-практиков, так и специалистов по компьютерному зрению и исследователей средств искусственного интеллекта.

В ближайшей перспективе, на пути повышения степени автономности РТК перед бортовой СТЗ стоят такие задачи как: определение общих условий движения РТК; автоматический поиск/выделение ориентиров в окружающем пространстве. При этом предполагается использование алгоритмов различной природы: традиционных (с пошаговым объяснением типа обработки и её результатов), нейросетевые. Кроме того, в развитие технологии сигнатур, эти структуры и механизмы работы с ними предлагается дополнить понятиями нечёткой логики и средствами мягких вычислений, что позволит получать информацию/принимать решения в ситуациях не полностью определённых и неоднозначных.

#### БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Minsky M., Papert S.* Perceptrons. – Cambridge, Mass.: MIT Press, 1969.
2. *Дуда П., Харт П.* Распознавание образов и анализ сцен. – М.: Мир, 1976.
3. *Liang-Chieh Chen, Yukun Zhu, Papandreou G., Schroff Florian, Hartwig Adam.* Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation, ECCV, 2018, Corpus ID: 3638670.
4. *Salas-Moreno R., Newcombe R., Strasdat H., Kelly P., and Davison A.* SLAM++: Simultaneous Localisation and Mapping at the Level of Objects // In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – IEEE, 2013. – P. 1352-1359.
5. *Newcombe R.A., Lovegrove S.J., and Davison A.J.* DTAM: Dense Tracking and Mapping in Real-Time // In Proceedings of the IEEE International Conference on Computer Vision (ICCV). – IEEE, 2011. – P. 2320-2327.

6. Bao S.Y., Bagra M., Chao Y.W., and Savarese S. Semantic structure from motion with points, regions, and objects // In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – IEEE, 2012. – P. 2703-2710.
7. Соколов С.М. Сравнительный анализ степени автономности робототехнических комплексов // Известия ЮФУ. Технические науки. – 2023. – № 1 (231). – С. 65-76. – <http://izv-tn.tti.sfedu.ru>.
8. Nüchter A. and Hertzberg J. Towards semantic maps for mobile robots // Robotics and Autonomous Systems. – 2008. – Vol. 56, No. 11. – P. 915-926.
9. Galindo C., Gonzalez J., Fernandez-Madriral J.A. Interactive In-Vehicle Guidance through a Multihierarchical Representation of Urban Maps // International journal of intelligent systems. – Vol. 25. – P. 597-620.
10. Ruijiao Li, Lai Wei, Dongbing Gu, Huosheng Hu, Klaus D. McDonald-Maier Multi-layered Map based Navigation and Interaction for an Intelligent Wheelchair // Proceeding of the IEEE International Conference on Robotics and Biomimetics (ROBIO) Shenzhen, China, December 2013. – P. 115-120.
11. Galindo C., Saffiotti A., Coradeschi S., Buschka P., Fernandez-Madriral J., and Gonzalez J. Multi-hierarchical semantic maps for mobile robotics // in Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Edmonton, CA, 2005. – P. 3492-3497. – Online at <http://www.aass.oru.se/asaffio/>.
12. Martinez Mozos O. and Burgard W. Supervised learning of topological maps using semantic information extracted from range data // In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Beijing, China, 2006. – P. 2772-2777.
13. Goerke N. and Braun S. Building semantic annotated maps by mobile robots, in Proceedings of the Conference Towards Autonomous Robotic Systems, Londonderry, UK, 2009.
14. Brunskill E., Kollar T., and Roy N. Topological mapping using spectral clustering and classification // in Proc. of IEEE/RSJ Conference on Robots and Systems (IROS). – 2007.
15. Friedman S., Pasula H., and Fox D. Voronoi random fields: Extracting the topological structure of indoor environments via place labeling // in Proc. of 19th International Joint Conference on Artificial Intelligence (IJCAI). – 2007.
16. Mozos O.M., Mizutani H., Kurazume R., and Hasegawa T. Categorization of indoor places using the kinect sensor // Sensors. – 2012. – Vol. 12, No. 5. – P. 6695-6711.
17. Diosi A., Taylor G., and Kleeman L. Interactive slam using laser and advanced sonar // in Proceedings of the IEEE International Conference on Robotics and Automation, Barcelona, Spain, 2005. – P. 1103-1108. Gard, “Conceptual spatial representations for indoor mobile robots,” Robotics and Autonomous Systems. – 2008. – Vol. 56, No. 6. – P. 493-502.
18. Zender H., Martinez Mozos O., Jensfelt P., Kruijff G., and Bur-Annual W. Conference on Human-Robot Interaction (HRI'06), Salt Lake.
19. Pronobi A. s and Jensfelt P. Multi-modal semantic mapping, in The RSS'11 Workshop on Grounding Human-Robot Dialog for Spatial Tasks, Los Angeles, CA, USA, July 2011. [Online]. Available: <http://www.pronobis.pro/publications/pronobis2011rss-ghrdst>.
20. Nieto-Granda C., III J.G.R., Trevor A.J.B., and Christensen H.I. Semantic map partitioning in indoor environments using regional analysis // in 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems, October 18-22, 2010, Taipei, Taiwan. IEEE, 2010. – P. 1451-1456.
21. Sabrina Wilske & Geert-Jan Kruijff Service Robots dealing with indirect speech acts. Language Technology Lab German Research Center for Artificial Intelligence (DFKI) Saarbrücken, Germany.
22. Randell D., Cui Z., Cohn A. A spatial logic based on regions and connection // In: Proceedings of the 3rd. International Conference on Knowledge Representation and Reasoning, San Mateo, Morgan Kaufmann, 1992. – P. 165-176;
23. Cadena C., Carlone L., Carrillo H., Latif Y., Scaramuzza D., Neira J. and Reid I., Leonard J.J. Past, Present, and Future of Simultaneous Localization And Mapping: Towards the Robust-Perception Age // in IEEE Transactions on Robotics. – 32 (6). – P. 1309-1332.
24. Everingham M., Van-Gool L., Williams C.K.I., Winn J., and Zisserman A. The PASCAL Visual Object Classes (VOC) Challenge // International Journal of Computer Vision. – 2010. – 88 (2). – P. 303-338.
25. Newcombe R.A., Lovegrove S.J., and Davison A.J. DTAM: Dense Tracking and Mapping in Real-Time // In Proceedings of the IEEE International Conference on Computer Vision (ICCV). – IEEE, 2011. – P. 2320-2327.

26. *Flint A., Murray D., and Reid I.D.* Manhattan Scene Understanding Using Monocular, Stereo, and 3D Features // In Proceedings of the IEEE International Conference on Computer Vision (ICCV). – IEEE, 2011. – P. 2228-2235.
27. *Bao S.Y., Bagra M., Chao Y.W., and Savarese S.* Semantic structure from motion with points, regions, and objects // In Proceedings of the IEEE Conference on Computer Vis.
28. *Krishnan R.G., Shalit U., and Sontag D.* Deep Kalman Filters // In NIPS 2016 Workshop: Advances in Approximate Bayesian Inference. – NIPS, 2016. – P. 1-7.
29. *Michael Wray, Diane Larlus, Gabriela Csurka, Dima Damen.* Fine-Grained Action Retrieval Through Multiple Parts-of-Speech Embeddings // ICCV 2019 paper is the Open Access version, provided by the Computer Vision Foundation, is available on IEEE Xplore. – [https://openaccess.thecvf.com/content\\_ICCV\\_2019/papers/Wray\\_Fine-Grained\\_Action\\_Retrieval\\_Through\\_Multiple\\_Parts-of-Speech\\_Embeddings\\_ICCV\\_2019\\_paper.pdf](https://openaccess.thecvf.com/content_ICCV_2019/papers/Wray_Fine-Grained_Action_Retrieval_Through_Multiple_Parts-of-Speech_Embeddings_ICCV_2019_paper.pdf).
30. *Богуславский А.А., Боровин Г.К., Карташев В.А., Павловский В.Е., Соколов С.М.* Модели и алгоритмы для интеллектуальных систем управления. – М.: ИПМ им. М.В. Келдыша, 2019. – 228 с.
31. *Соколов С.М., Богуславский А.А., Беклемишев Н.Д.* Программное обеспечение систем технического зрения реального времени для систем управления робототехническими комплексами // Матер. XIII Международной научно-технической конференции «Завалишинские чтения – 2018». – СПб.: ГУАП, 16-20 апреля 2018 г. – С. 205-211.
32. *Соколов С.М., Богуславский А.А., Беклемишев Н.Д.* Реализация интерпретирующей навигации с помощью модулей СТЗ // Матер. 30-й Международной научно-технической конференции «Экстремальная робототехника», 13-15 июня 2019, Санкт-Петербург: ЦНИИ РТК. – С. 264-267.
33. *Thrun S. and Bucken A.* Integrating grid-based and topological maps for mobile robot navigation // In Proceedings of the National Conference on Artificial Intelligence. – 1996. – No. 8.
34. *Mozos O. and Burgard W.* Supervised Learning of Topological Maps using Semantic Information Extracted from Range Data // 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, – Oct. 2006. – P. 2772-2777.
35. *Pronobis A. and Jensfelt P.* Understanding the real world: Combining objects, appearance, geometry and topology for semantic mapping. Royal Institute of Technology (KTH) SE-100, Tech. Rep., 2011.
36. *Dansereau D.G., Williams S.B., and Corke P.* Simple Change Detection from Mobile Light Field Cameras. Computer Vision and Image Understanding. – 2016.
37. *Davison A., I. Reid, N. Molton, and O. Stasse.* MonoSLAM: RealTime Single Camera SLAM // IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI). – 2007. – 29 (6). – P. 1052-1067.
38. *Соколов С.М.* Онтологический подход в создании робототехнических комплексов с повышенной степенью автономности // Известия ЮФУ. Технические науки. – 2022. – № 1. – С. 42-59.
39. *Васильев С.Н., Жерлов А.К., Федосов Е.А., Федунев Е.Е.* Интеллектуальное управление динамическими системами. – М.: Физико-математическая литература, 2000. – 352 с.

## REFERENCES

1. *Minsky M., Papert S.* Perceptrons. Cambridge, Mass.: MIT Press, 1969.
2. *Duda R., Khart P.* Raspoznavanie obrazov i analiz stsen [Pattern recognition and scene analysis]. Moscow: Mir, 1976.
3. *Liang-Chieh Chen, Yukun Zhu, Papandreou G., Schroff Florian, Hartwig Adam.* Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation, ECCV, 2018, Corpus ID: 3638670.
4. *Salas-Moreno R., Newcombe R., Strasdat H., Kelly P., and Davison A.* SLAM++: Simultaneous Localisation and Mapping at the Level of Objects, In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2013, pp. 1352-1359.
5. *Newcombe R.A., Lovegrove S.J., and Davison A.J.* DTAM: Dense Tracking and Mapping in Real-Time, In Proceedings of the IEEE International Conference on Computer Vision (ICCV). IEEE, 2011, pp. 2320-2327.
6. *Bao S.Y., Bagra M., Chao Y.W., and Savarese S.* Semantic structure from motion with points, regions, and objects, In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2012, pp. 2703-2710.

7. Sokolov S.M. Sravnitel'nyy analiz stepeni avtonomnosti robototekhnicheskikh kompleksov [Comparative analysis of the degree of autonomy of robotic complexes], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2023, No. 1 (231), pp. 65-76. Available at: <http://izv-tn.tti.sfedu.ru>.
8. Nüchter A. and Hertzberg J. Towards semantic maps for mobile robots, *Robotics and Autonomous Systems*, 2008, Vol. 56, No. 11, pp. 915-926.
9. Galindo C., Gonzalez J., Fernandez-Madriral J.A. Interactive In-Vehicle Guidance through a Multihierarchical Representation of Urban Maps, *International journal of intelligent systems*, Vol. 25, pp. 597-620.
10. Ruijiao Li, Lai Wei, Dongbing Gu, Huosheng Hu, Klaus D. McDonald-Maier Multi-layered Map based Navigation and Interaction for an Intelligent Wheelchair, *Proceeding of the IEEE International Conference on Robotics and Biomimetics (ROBIO) Shenzhen, China, December 2013*, pp. 115-120.
11. Galindo C., Saffiotti A., Coradeschi S., Buschka P., Fernandez-Madriral J., and Gonzalez J. Multi-hierarchical semantic maps for mobile robotics, in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Edmonton, CA, 2005*, pp. 3492-3497. Online at <http://www.aass.oru.se/asaffio/>.
12. Martinez Mozos O. and Burgard W. Supervised learning of topological maps using semantic information extracted from range data, in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Beijing, China, 2006*, pp. 2772-2777.
13. Goerke N. and Braun S. Building semantic annotated maps by mobile robots, in *Proceedings of the Conference Towards Autonomous Robotic Systems, Londonderry, UK, 2009*.
14. Brunskill E., Kollar T., and Roy N. Topological mapping using spectral clustering and classification, in *Proc. of IEEE/RSJ Conference on Robots and Systems (IROS), 2007*.
15. Friedman S., Pasula H., and Fox D. Voronoi random fields: Extracting the topological structure of indoor environments via place labeling, in *Proc. of 19th International Joint Conference on Artificial Intelligence (IJCAI), 2007*.
16. Mozos O.M., Mizutani H., Kurazume R., and Hasegawa T. Categorization of indoor places using the kinect sensor, *Sensors*, 2012, Vol. 12, No. 5, pp. 6695-6711.
17. Diosi A., Taylor G., and Kleeman L. Interactive slam using laser and advanced sonar, in *Proceedings of the IEEE International Conference on Robotics and Automation, Barcelona, Spain, 2005*, pp. 1103-1108. Gard, Conceptual spatial representations for indoor mobile robots // *Robotics and Autonomous Systems*, 2008, Vol. 56, No. 6, pp. 493-502.
18. Zender H., Martinez Mozos O., Jensfelt P., Kruijff G., and Burdick W. Conference on Human-Robot Interaction (HRI'06), Salt Lake.
19. Pronobis A. and Jensfelt P. Multi-modal semantic mapping, in *The RSS'11 Workshop on Grounding Human-Robot Dialog for Spatial Tasks, Los Angeles, CA, USA, July 2011*. [Online]. Available: <http://www.pronobis.pro/publications/pronobis2011rss-ghrdst>.
20. Nieto-Granda C., III J.G.R., Trevor A.J.B., and Christensen H.I. Semantic map partitioning in indoor environments using regional analysis, in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems, October 18-22, 2010, Taipei, Taiwan. IEEE, 2010*, pp. 1451-1456.
21. Sabrina Wilske & Geert-Jan Kruijff Service Robots dealing with indirect speech acts. Language Technology Lab German Research Center for Artificial Intelligence (DFKI) Saarbrücken, Germany.
22. Randell D., Cui Z., Cohn A. A spatial logic based on regions and connection, in: *Proceedings of the 3rd. International Conference on Knowledge Representation and Reasoning, San Mateo, Morgan Kaufmann, 1992*, pp. 165-176;
23. Cadena C., Carlone L., Carrillo H., Latif Y., Scaramuzza D., Neira J. and Reid I., Leonard J.J. Past, Present, and Future of Simultaneous Localization And Mapping: Towards the Robust-Perception Age, in *IEEE Transactions on Robotics*, 32 (6), pp. 1309-1332.
24. Everingham M., Van-Gool L., Williams C K.I., Winn J., and Zisserman A. The PASCAL Visual Object Classes (VOC) Challenge, *International Journal of Computer Vision*, 2010, 88 (2), pp. 303-338.
25. Newcombe R.A., Lovegrove S.J., and Davison A.J. DTAM: Dense Tracking and Mapping in Real-Time, in *Proceedings of the IEEE International Conference on Computer Vision (ICCV). IEEE, 2011*, pp. 2320-2327.

26. Flint A., Murray D., and Reid I.D. Manhattan Scene Understanding Using Monocular, Stereo, and 3D Features, *In Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. IEEE, 2011, pp. 2228-2235.
27. Bao S.Y., Bagra M., Chao Y.W., and Savarese S. Semantic structure from motion with points, regions, and objects, *In Proceedings of the IEEE Conference on Computer Vis.*
28. Krishnan R.G., Shalit U., and Sontag D. Deep Kalman Filters, *In NIPS 2016 Workshop: Advances in Approximate Bayesian Inference*. NIPS, 2016, pp. 1-7.
29. Michael Wray, Diane Larlus, Gabriela Csurka, Dima Damen. Fine-Grained Action Retrieval Through Multiple Parts-of-Speech Embeddings, *ICCV 2019 paper is the Open Access version, provided by the Computer Vision Foundation, is available on IEEE Xplore*. Available at: [https://openaccess.thecvf.com/content\\_ICCV\\_2019/papers/Wray\\_Fine-Grained\\_Action\\_Retrieval\\_Through\\_Multiple\\_Parts-of-Speech\\_Embeddings\\_ICCV\\_2019\\_paper.pdf](https://openaccess.thecvf.com/content_ICCV_2019/papers/Wray_Fine-Grained_Action_Retrieval_Through_Multiple_Parts-of-Speech_Embeddings_ICCV_2019_paper.pdf).
30. Boguslavskiy A.A., Borovin G.K., Kartashev V.A., Pavlovskiy V.E., Sokolov S.M. Modeli i algoritmy dlya intellektual'nykh sistem upravleniya [Models and algorithms for intelligent control systems]. Moscow: IPM im. M.V. Keldysha, 2019, 228 p.
31. Sokolov S.M., Boguslavskiy A.A., Beklemishev N.D. Programmnoe obespechenie sistem tekhnicheskogo zreniya real'nogo vremeni dlya sistem upravleniya robototekhnicheskimi kompleksami [Software for real-time vision systems for control systems of robotic complexes], *Mater. XIII Mezhdunarodnoy nauchno-tekhnicheskoy konferentsii «Zavalishinskie chteniya – 2018»* [Materials of the XIII International Scientific and Technical Conference "Zavalishinsky Readings – 2018"]. St. Petersburg: GUAP, 16-20 aprelya 2018 g., pp. 205-211.
32. Sokolov S.M., Boguslavskiy A.A., Beklemishev N.D. Realizatsiya interpretiruyushchey navigatsii s pomoshch'yu moduley STZ [Implementation of interpretive navigation using STZ modules], *Mater. 30-y Mezhdunarodnoy nauchno-tekhnicheskoy konferentsii «Ekstremal'naya robototekhnika», 13-15 iyunya 2019, Sankt-Peterburg: TSNII RTK* [Proceedings of the 30th International Scientific and Technical Conference "Extreme Robotics", June 13-15, 2019, St. Petersburg: Central Research Institute of RTK], pp. 264-267.
33. Thrun S. and Bucken A. Integrating grid-based and topological maps for mobile robot navigation, *In Proceedings of the National Conference on Artificial Intelligence*, 1996, No. 8.
34. Mozos O. and Burgard W. Supervised Learning of Topological Maps using Semantic Information Extracted from Range Data, *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct. 2006, pp. 2772-2777.
35. Pronobis A. and Jensfelt P. Understanding the real world: Combining objects, appearance, geometry and topology for semantic mapping. Royal Institute of Technology (KTH) SE-100, Tech. Rep., 2011.
36. Dansereau D.G., Williams S.B., and Corke P. Simple Change Detection from Mobile Light Field Cameras. *Computer Vision and Image Understanding*, 2016.
37. Davison A., I. Reid, N. Molton, and O. Stasse. MonoSLAM: RealTime Single Camera SLAM, *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 2007, 29 (6), pp. 1052-1067.
38. Sokolov S.M. Ontologicheskii podkhod v sozdanii robototekhnicheskikh kompleksov s povyshennoy stepen'yu avtonomnosti [An ontological approach to the creation of robotic complexes with an increased degree of autonomy], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2022, No. 1, pp. 42-59.
39. Vasil'ev S.N., Zherlov A.K., Fedosov E.A., Fedunov E.E. Ingellektnoe upravlenie dinamicheskimi sistemami [Intelligent control of dynamic systems]. Moscow: Fiziko-matematicheskaya literatura, 2000, 352 p.

Статью рекомендовал к опубликованию д.т.н. М.Ю. Медведев.

**Соколов Сергей Михайлович** – Институт прикладной математики им. М.В. Келдыша РАН; e-mail: sokolsm@keldysh.ru; г. Москва, Россия; тел.: 84992207994; д.ф.-м.н.; профессор; г.н.с.; зав. сектором.

**Sokolov Sergey Mikhailovich** – Keldysh Institute of Applied Mathematics RAS; e-mail: sokolsm@keldysh.ru; Moscow, Russia; phone: +74992207994; dr. of math. and physics. sc.; professor; chief researcher; chief division.