

На правах рукописи



Вохминцев Александр Владиславович

**МЕТОДОЛОГИЯ РЕШЕНИЯ ПРОБЛЕМЫ ОДНОВРЕМЕННОЙ
НАВИГАЦИИ И ПОСТРОЕНИЯ КАРТЫ НА ОСНОВЕ
КОМБИНИРОВАНИЯ ВИЗУАЛЬНЫХ И СЕМАНТИЧЕСКИХ
ХАРАКТЕРИСТИК ОКРУЖАЮЩЕЙ СРЕДЫ**

Специальность:

**05.13.01 – Системный анализ, управление и обработка информа-
ции (в информационных и технических системах)**

**Автореферат диссертации на соискание ученой степени
доктора технических наук**

Челябинск – 2020

Работа выполнена в Челябинском государственном университете.

Научный консультант: доктор технических наук, профессор, **Мельников Андрей Витальевич**

Официальные оппоненты:

Мещеряков Роман Валерьевич, доктор технических наук, профессор РАН, ФГБУН «Институт проблем управления им. В.А. Трапезникова РАН», главный научный сотрудник

Даринцев Олег Владимирович, доктор технических наук, профессор, ФГБОУ ВО «Уфимский государственный авиационный технический университет», профессор кафедры автоматизации технологических процессов

Тебуева Фариза Биляловна, доктор физико-математических наук, доцент, ФГАОУ ВО «Северо-Кавказский федеральный университет», заведующий кафедрой прикладной математики и компьютерной безопасности

Ведущая организация: Федеральное государственное учреждение «Федеральный исследовательский центр «Информатика и управление» Российской академии наук» (ФИЦ ИУ РАН), г. Москва

Защита диссертации состоится 15 декабря 2020 г. в 10⁰⁰ часов на заседании диссертационного совета Д 212.288.12 на базе ФГБОУ ВО «Уфимский государственный авиационный технический университет» по адресу: 450008, г. Уфа, ул. К. Маркса, 12.

С диссертацией можно ознакомиться в библиотеке ФГБОУ ВО «Уфимский государственный авиационный технический университет» и на сайте www.ugatu.su.

Автореферат разослан «___» _____ 2020 года.

Ученый секретарь
диссертационного совета



Сметанина Ольга Николаевна

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. Проблема одновременной навигации и построения карты (Simultaneous localization and mapping, SLAM) заключается в реконструкции двумерной или трехмерной карты неизвестного окружающего пространства вокруг автономного интеллектуального объекта управления, например, мобильной платформы и определении своего положения на этой карте в каждый момент времени во время движения. Проблема SLAM является одной из ключевых проблем в современной робототехнике и машинном зрении, так как на ее решении основано создание автономных интеллектуальных робототехнических комплексов. Данная проблема считается полностью решенной для построения двумерных карт и трехмерных карт для небольших статических сцен, однако для динамических, контекстуально сложных и крупномасштабных сцен является активной областью научных исследований. По оценкам аналитиков агентства по перспективным оборонным научно-исследовательским разработкам (DARPA) существуют следующие актуальные проблемы в области SLAM: разработка быстрых методов для навигации в неизвестном динамическом пространстве; разработка интеллектуальных методов навигации на основе семантических свойств окружающей среды; создание методов для навигации на сценах, искаженных помехами окружающей среды различной природы; проблема позиционной точности методов построения трехмерной карты, связанная с принципиальной возможностью построения высоко точных моделей трехмерных объектов на основе визуальной информации, одометрической информации и информации о глубине. Данное исследование направлено на решение второй проблемы. Решение представленных проблем в области SLAM позволит перейти к созданию новых видов роботизированных коллективных систем, с помощью которых будет решаться широкий спектр задач, таких как: распознавание и слежение за динамическими объектами в заданном пространстве, предсказание траектории их перемещения, согласованное и оптимизированное планирование движений мобильных платформ в заданном пространстве с учетом возможных статических и динамических препятствий. Важным направлением применения результатов диссертационного исследования является коллаборативная робототехника, основной тренд в которой связан с созданием комплексов, оснащенных собственной системой машинного зрения и системой SLAM. Перспективным направлением применения результатов исследования является антропоморфная робототехника и связанные с ней области исследований.

Степень разработанности проблемы. Принято различать следующие методологии SLAM: визуальная (Visual SLAM), на основе частичного фильтра (Particle Filter SLAM, FAST SLAM), на основе расширенного фильтра Калмана (Extended Kalman Filter, EKF), на основе графов (Graph-Based SLAM). Известным методологиям свойственны следующие недостатки:

– методологии имеют ограничения по применению в окружающей среде, искаженной различными естественными помехами, в том числе классическая методология Visual SLAM показывает неудовлетворительные результаты для сцен с неравномерным освещением;

- методологии характеризуются высокой вычислительной и алгоритмической сложностью, поэтому их применение ограничено в приложениях, где требуется осуществлять одновременную навигацию и картографирование для контекстуально сложных крупномасштабных сцен в режиме реального времени;
- методология EKF-SLAM не учитывает при решении поставленной задачи визуальную информацию об окружающей среде. Оценка карты и определение положения навигационных ориентиров на сцене осуществляются без использования данных о форме реконструируемой поверхности объекта, поэтому точность получаемых оценок положения мобильной платформы на карте относительно навигационных ориентиров не высока;
- в методологии Visual SLAM информация о визуально связанных характеристиках и трехмерных облаках точек не используется для совместного (комбинированного) решения вариационной задачи ICP. Результат реконструкции сцены с точки зрения точности и сходимости обусловлен выбором начального приближения для решения вариационной задачи;
- визуальные методы для решения проблемы SLAM, которые основаны на итеративном алгоритме ICP для группы аффинных преобразований, не позволяют проводить точную регистрацию структурных элементов сцены;
- в методологии Visual SLAM не используется информация о семантических характеристиках окружающей среды, в следствие чего все особые точки в двумерных и трехмерных данных имеют одинаковую значимость;
- в методологии Visual SLAM информация о визуально связанных характеристиках окружающей среды используется для улучшения точности и сходимости всей вычислительной процедуры только на этапе сопоставления данных об особых точках на изображении.

Для преодоления указанных недостатков в работе предлагается разработать новую методологию решения проблемы SLAM на основе комбинирования визуальных и семантических характеристик окружающей среды. Предложенная методология основана на использовании трех типов данных: визуально связанные характеристики окружающей среды (изображения, данные о цвете), трехмерные облака точек (данные о глубине сцены), данные о семантических свойствах окружающей среды, при этом визуальная и семантическая информация о трехмерной сцене будет использоваться на всех ключевых шагах решения проблемы SLAM.

Проблемы создания теории, моделей, методов и алгоритмов для создания системы одновременной навигации и картографирования рассматриваются в следующих научных трудах: T. Rui, Q. Zhang, Y. Zhou, J. Xing, G. Dissanayake, P. Newman, S. Clark, H.-F. Durrant-Whyte, M. Csorba, S.K. Zhou, R. Chellappa, B. Moghaddam, M. Montemerlo, S. Thrun, D. Koller, B. Wegbreit, C. Estrada, J. Neira, J.D. Tardós, В.И. Кобера, Е.Н. Соснова, С.Л. Зенкевича и В.С. Лемпицкого. Было создано ряд передовых технологий в области SLAM в неизвестной среде, среди которых можно выделить разработки Stanford Artificial Intelligence Lab, Stanford Racing Team, Volkswagen Electronics Research Laboratory, Massachusetts Institute

of Technology Lincoln Laboratory, Pixon Imaging. Данные технологии нашли применение в военной сфере для организации наблюдения на поле боя или при разведывательных операциях, для обеспечения безопасности военных и стратегических промышленных объектов, при организации спасательных операций и для эвакуации раненных с поля боя. В работе рассмотрено применение предлагаемой комбинированной методологии SLAM для создания систем спасения людей. Диссертационная работа связана с тематикой работ, осуществляемых в соответствии с Национальным проектом «Цифровая экономика». На основании сказанного выше сформулирована основная цель диссертационной работы и задачи исследования.

Целью диссертационной работы является повышение эффективности процессов обработки информации в методах и алгоритмах для реконструкции трехмерных сцен окружающего пространства, визуализации трёхмерных моделей и навигации мобильных платформ в неизвестной среде на основе комбинирования информации о семантических характеристиках окружающей среды и доступной мультисенсорной информации, получаемой с различных типов датчиков. Для достижения поставленной цели в диссертационной работе решаются следующие задачи:

1. Разработать методологию для точного решения задачи SLAM, основанную на комбинировании визуальных и семантических характеристик неизвестной окружающей среды (соответствует п. 2 паспорта специальности).

2. Разработать алгоритм поиска и сопоставления особых точек на основе двумерных изображений для обработки информации о цвете в кадре (соответствует п. 4 паспорта специальности).

3. Разработать точные комбинированные методы реконструкции контекстуально сложных крупномасштабных трехмерных сцен на основе итеративного алгоритма ближайших точек для группы аффинных преобразований и для группы ортогональных преобразований (соответствует п. 4 и п. 12 паспорта специальности).

4. Разработать метод решения задачи навигации и определения положения мобильной платформы на сцене, основанный на комбинировании визуальных и семантических характеристик неизвестной окружающей среды и расширенной Калмановской фильтрации (соответствует п. 4 и п. 6 паспорта специальности).

5. Разработать быстрый рекурсивный алгоритм для вычисления дискретного косинусного преобразования (ДКП) для решения задач сжатия и восстановления изображений, получаемых с множества датчиков на мобильной платформе (соответствует п. 4 паспорта специальности).

6. Предложить подход к оптимизации графа положений мобильной платформы и обнаружению «замыканий цикла» на основе метода корзина слов и комбинировании визуальных и семантических характеристик неизвестной окружающей среды (соответствует п. 4 паспорта специальности).

7. Произвести апробацию предложенной комбинированной визуальной системы SLAM для решения задачи спасения людей (соответствует п. 5 паспорта специальности).

Объектом исследования диссертационной работы является система обработки визуальной и семантической информации о неизвестной окружающей среде для решения задачи реконструкции контекстуально сложных крупномасштабных трехмерных сцен и задачи определения положения мобильной платформы на сцене в каждый момент времени.

Предметом исследования диссертационной работы являются методы, алгоритмы для решения проблемы одновременной навигации и построения карты на основе комбинирования информации о семантических и визуальных характеристиках окружающей среды.

Методы исследования. Повышение степени автономности управления роботизированным комплексом – основная цель, преследуемая исследователями при разработке современных робототехнических систем. Достижение этой цели возможно за счет интеграции теории автоматического управления, теории искусственного интеллекта, компьютерного когнитивного моделирования и создания на основе такой интеграции систем одновременной навигации и построения карты нового поколения.

Научная новизна диссертационной работы заключается в следующем:

1. Предложена новая методология для решения проблемы одновременной навигации и построения карты, основанная на комбинировании визуальных и семантических характеристик неизвестной окружающей среды. Разработанная методология позволяет осуществлять реконструкцию трехмерной карты неизвестной окружающей среды с использованием семантических характеристик объектов на сцене и доступной мультисенсорной информации, поступающей с датчиков при произвольном изменении их положения в трехмерном пространстве. Предложенная методология основана на визуальной методологии SLAM и отличается от известных методологий тем, что:

– информация о визуально связанных характеристиках окружающей среды используется на всех этапах предлагаемой методологии для улучшения точности и сходимости соответствующих предложенных методов и алгоритмов, а не только на этапе сопоставления данных об особых точках на изображении;

– информация о семантических характеристиках окружающей среды извлекается из данных о цвете на основе методов семантической маркировки и сегментации пространства, при этом принадлежность к определенному классу предметной области влияет на значение весовых коэффициентов, которые связаны с семантической маркировкой окружающего пространства;

– информация о визуально связанных характеристиках, трехмерных облаках точек, информация о семантических характеристиках окружающей среды используется для совместного (комбинированного) решения вариационной задачи ИСР;

– информация о семантических характеристиках окружающей среды используется при решении вариационной задача точка-плоскость для класса ортогональных преобразований, при решении вариационной задача точка-точка для

класса аффинных преобразований, при определении положения мобильной платформы в неизвестном пространстве: семантические маркеры учитываются при расчете матрицы перехода состояний и вектора управлений;

- позволяет решать проблему SLAM для контекстуально сложных крупномасштабных сцен в реальном масштабе времени в контролируемых и неконтролируемых условиях, связанных с условиями равномерного или неравномерного освещения на сцене.

2. Разработан алгоритм поиска особых точек на изображениях в последовательности кадров на основе дескриптора локальной гистограммы ориентированных градиентов (ДЛ ГНГ) с рекурсивной процедурой вычисления, который отличается тем, что:

- обладает лучшими характеристиками по точности в сравнении с известными дескрипторами при малых поворотах вне области сцены;

- используется для сопоставления изображений с произвольным пространственным разрешением и небольшим изменением масштаба относительно друг друга, дает точные оценки для сложных крупномасштабных сцен;

- может применяться в приложениях при сопоставлении изображений для крупномасштабных сцен в реальном режиме времени за счет использования процедуры дискретизации изображения в виде пирамидального разложения на несколько связанных квантованных уровней и процедуры предварительной обработки данных, связанной с вычислением значения нормализованных центрированных ГНГ и преобразований Фурье.

3. Разработаны комбинированные методы для реконструкции контекстуально сложных крупномасштабных трехмерных сцен для группы аффинных и для группы ортогональных преобразований на основе решения вариационной задачи ИСР в замкнутой форме с использованием метрики точка-точка и точка-плоскость соответственно с использованием визуально связанных и семантических характеристик окружающей среды. Предложенные комбинированные методы улучшают качество работы двух ключевых шагов метода ИСР: определение соответствующих точек между парой трехмерных облаков точек и решение вариационной задачи и отличаются тем, что:

- позволяют решить проблему зависимости результата решения вариационной задачи от правильности выбора начальных значений;

- используются для регистрации трехмерных облаков точек с произвольным пространственным разрешением и масштабом относительно друг друга и дают точные оценки для сложных крупномасштабных сцен;

- применение визуально связанных характеристик на основе предложенного дескриптора (ДЛ ГНГ) для решения вариационной задачи алгоритма ИСР для аффинных преобразований позволяет получить в неконтролируемых условиях лучшую сходимость и точность в сравнении с другими дескрипторными методами: метод регистрации сходится уже после 11 итерации, тогда как при использовании дескриптора ORB предложенный метод сходится только после 16 итерации;

- применение визуально связанных характеристик на основе предложенного дескриптора (ДЛ ГНГ) для решения вариационной задачи алгоритма ИСР для ортогональных преобразований позволяет получить в неконтролируемых

условиях лучшую сходимость и точность в сравнении с другими дескрипторными методами: предложенный метод регистрации сходиться уже после 10 итерации, тогда как при использовании дескриптора ORB предложенный метод регистрации сходиться только после 14 итерации;

- предложенный метод решения вариационной задачи для ортогональных преобразований в контролируемых условиях имеет сходимость близкую с методом, который использует метрику точка-точка с экстраполяцией, а в неконтролируемых условиях предложенный метод сходится на 1-2 итерацию позже, чем метод регистрации данных с метрикой точкой-плоскость, но при этом показывает лучшую точность;

- решение для группы аффинных преобразований позволяет проводить точную регистрацию структурных элементов сцены (вырожденные случаи) и создает основу для применения метода Хорна на случай с не ригидными объектами на сцене.

В работе установлено, что точность реконструкции зависит от количества особых точек в кадре нелинейным образом – в виде функции с одним ярко выраженным пиком для всех типов дескрипторов.

4. Предложен метод решения задачи навигации и определения положения мобильной платформы на сцене (S-EKF SLAM), основанный на комбинировании визуальных и семантических характеристик неизвестной окружающей среды и расширенной Калмановской фильтрации (EKF), получены формулы для оценки вектора состояния при движении мобильной платформы по серии зашумленных измерений и семантических меток. Метод S-EKF SLAM отличается тем, что:

- комбинирование информации происходит на уровне выделенных семантических и мультимодальных визуальных характеристик;

- в контролируемых условиях метод с точки зрения точности выдает результаты подобные Visual SLAM, но при этом предложенный метод имеет лучшую сходимость;

- в неконтролируемых условиях метод с точки зрения точности выдает лучшие результаты, чем методы определения положения мобильной платформы в известных методологиях Visual SLAM, EKF SLAM, Graph-SLAM;

- имеет хорошую сходимость и может использоваться в приложениях, работающих в реальном масштабе времени. Сходимость предложенного метода улучшается за счет использования семантических ориентиров в структуре расширенного фильтра Калмана.

5. Разработан быстрый рекурсивный алгоритм для вычисления ДКП дискретного сигнала в скользящем окне с произвольным шагом. Алгоритм основан на рекурсивном соотношении между тремя последовательными локальными спектрами ДКП, которые вычисляются в равноудаленных положениях окна, который отличается тем, что является более эффективным, чем известные быстрые алгоритмы ДКП, когда шаг скользящего окна меньше, чем граничное значение, равное 8. Предложенный алгоритм используется для сжатия изображений, получаемых с множества датчиков, установленных на мобильной платформе.

6. Предложен подход к оптимизации графа положений мобильной платформы и обнаружению «замыканий цикла» с использованием метода корзина слов и предложенного двумерного дескриптора (ДЛ ГНГ), который отличается тем, что:

- позволяет проводить корректировку движения мобильной платформы и уточнения трехмерной карты окружающей среды, при этом точность локализации для коротких циклов увеличивается в пределах диапазона от 15% до 25% для разных серий тестов, тогда как для длинных циклов различие в точности локализации может существенно отличаться (в несколько раз);

- использование предложенного дескриптора (ДЛ ГНГ) вместо известных дескрипторов (SUFT, SURF, ORB) позволяет значительно сократить вычислительную сложность этапа, связанного с определением визуального подобия изображений в двух ключевых последовательных кадрах данных;

- предложенная эффективная процедура поиска изображений по визуальному подобию на основе инвертированного индекса позволяет использовать подход на основе корзины слов в реальном масштабе времени.

Теоретическая и практическая значимость результатов работы.

Научную значимость исследования представляет методология решения проблемы одновременной навигации и трехмерного картографирования на основе комбинирования визуальных и семантических характеристик окружающей среды. Теоретическую значимость имеют следующие научные результаты в предложенной методологии: алгоритм поиска особых точек на изображениях; комбинированные методы решения вариационной задачи в замкнутой форме; метод решения задачи навигации и определения положения мобильной платформы; метод сжатия изображений; подход к оптимизации графа положений мобильной платформы и обнаружению «замыканий цикла». Практическую значимость работы составляют прототип комбинированной системы SLAM для решения задачи спасения людей и результаты экспериментальных исследований и компьютерного моделирования. Созданная на основе предложенных методов система может быть применена в системах: спасения людей; в некооперативных системах биометрической идентификации личности и интеллектуального наблюдения; автономной дистанционной навигации робототехнических комплексов в неизвестном пространстве, в том числе системах автономной навигации автомобилей, подводных или летательных аппаратов без участия человека; мониторинга за окружающим пространством в опасных или труднодоступных для человека объектах (зоны чрезвычайных ситуаций с разрушениями или поражением вредными веществами или излучениями для человека); с коллаборативной робототехникой в промышленном производстве и логистике. Исследование направлено на реализацию приоритетных направлений развития РФ в части перехода к роботизированным системам, создания систем машинного обучения и искусственного интеллекта, в том числе в сфере обеспечения общественной безопасности и противодействия терроризму.

Внедрение результатов исследований. Результаты внедрены на предприятиях: АО «Научно-производственное объединение Андроидная техника»,

УРТЦ «Альфа-Интех», ООО «Научно-производственный центр Интеллектуальные транспортные системы», в Югорском научно-исследовательском институте информационных технологий, в учебном процессе ФГБОУ ВО «Челябинский государственный университет». Результаты работы используются в учебном процессе Челябинского государственного университета, а именно в 2018 г. в университете создана магистерская программа «Робототехника» для направления 02.04.02 «Фундаментальная информатика и информационные технологии» в рамках выполнения гранта Благотворительного фонда В. Потанина в 2017-2018.

Апробация результатов работы. Основные положения и результаты, полученные в диссертационной работе, представлены и обсуждены на следующих научных мероприятиях: 5-ая Всероссийская научная конференция «Системный анализ и информационные технологии», Россия, Абзаково, 2011; 13-th International Workshop on “Computer Science and Information Technology”, Germany, Garmisch-Partenkirchen, 2011; XVII Всероссийская научная конференция «Информационные и математические технологии в науке и управлении», Россия, Иркутск, 2012; “New Prolamat”, Germany, Dresden, 2013; 15-th International Workshop on “Computer Science and Information Technology”, Austria, Vienna, 2013; XV Международная конференция «Проблемы управления и моделирования в сложных системах», Россия, Самара, 2013; 16-th International Workshop on “Computer Science and Information Technology”, United Kingdom, Sheffield, 2014; 3-ая Международная научная конференция «Информационные технологии и системы», Россия, Челябинск-Банное, 2014; XVI Международная конференция «Проблемы управления и моделирования в сложных системах», Россия, Самара, 2014; 6-ая Всероссийская научная конференция «Системный анализ и информационные технологии», Россия, Калининград, 2015; 4-ая Международная научная конференция «Информационные технологии и системы», Россия, Челябинск-Банное, 2015; 4-ая Международная научная конференция «Анализ изображений социальных сетей и текстов» (Analysis of images, social networks and texts), Россия, Екатеринбург, 2015; 17-th International Workshop on “Computer Science and Information Technology”, Italy, Rome, 2017; 1-ая Международная научно-техническая конференция «Пром-Инжиниринг» (IEEE International Conference on Industrial Engineering), Россия, Челябинск, 2015; SPIE Conference on “Applications of Digital Image Processing” XXXVIII, USA, San-Diego, 2015; 5-ая Международная научная конференция «Анализ изображений социальных сетей и текстов» (Analysis of images, social networks and texts), Россия, Екатеринбург, 2016; 2-ая Международная научно-техническая конференция «Пром-Инжиниринг» (IEEE International Conference on Industrial Engineering), Россия, Челябинск, 2016; SPIE Conference on “Applications of Digital Image Processing” XXXIX, USA, San-Diego, 2016; SPIE Conference on “Applications of Digital Image Processing” XL, USA San-Diego, 2017; 19-th International Workshop on “Computer Science and Information Technology”, Germany, Karlsruhe, 2017; 3-nd International Conference on “Industrial Engineering, Applications and Manufacturing” (IEEE ICIEAM) Russia, Saint-Petersburg, 2017; SPIE Conference on “Applications of Digital Image Processing” XLI, USA, San-Diego, USA, 2018; 20-th International workshop on “Computer science and information technologies”,

Bulgaria, Varna, 2018; 18-th IFAC Conference on “Technology, Culture and International Stability”, Azerbaijan, Baku, 2018; 7-ая Международная научная конференция «Информационные технологии и системы», Россия, Ханты-Мансийск, 2019.

Научные исследования в рамках диссертационной работы поддержаны грантами следующих научных фондов: РФФИ 16-08-00342 «Разработка системы одновременной навигации и составления карты в неизвестном пространстве на основе динамического сопоставления изображений и регистрации облаков точек», 2016-2018 гг. (руководитель Вохминцев А.В.); РФФИ 15-19-10010 «Разработка алгоритмической модели технической системы для идентификации личности по мультисенсорным биометрическим данным», 2015-2017 гг. (основной исполнитель, руководитель Кобер В.И.); РФФИ: 18-37-20032 «Методы управления автономными техническими объектами на основе планирования траектории в среде с динамическими препятствиями», 2018-2019 гг. (основной исполнитель, руководитель Яковлев К.С.); Государственное задание (проектная часть) Министерство образования и науки «Разработка адаптивных методов для надежного слежения за трехмерными объектами» 2.1766.2014/К, 2014-2016 гг. (основной исполнитель, руководитель Кобер В.И.); Государственное задание (проектная часть) Министерство образования и науки «Разработка адаптивных методов трехмерной реконструкции окружающего пространства по динамическим мультисенсорным данным» 2.1743.2017.ПЧ, 2017-2019 гг. (основной исполнитель, руководитель Кобер В.И.).

Обоснованность и достоверность результатов. Обоснованность полученных в диссертационной работе результатов основана на использовании апробированных методов исследования, корректном применении математического аппарата теории искусственного интеллекта, теории управления и методов обработки визуальных данных, согласовании полученных результатов с известными теоретическими положениями в области SLAM и трехмерной реконструкции. Адекватность и достоверность предложенных методов и алгоритмов подтверждается результатами компьютерного моделирования для эталонных баз данных, а также экспериментами с реальными данными. Точность и сходимость предложенных методов была оценена на эталонных базах данных NYU Depth Dataset и ASL DataSet. Был проведен сравнительный анализ методов SLAM в серии экспериментов в терминах точности и вычислительной сложности с известными открытыми системами SLAM HOGMap, ORB-SLAM, TORO-SLAM, RGBDSlam, EKFMonoSLAM, используя мобильные платформы Odyssey 6 Robotics и Husky UGV.

Публикации. По теме диссертационной работы опубликована 31 научная публикация: из них 12 научных статей опубликованы в ведущих российских научных журналах из списка ВАК, всего изданий из списка ВАК – 17; 14 публикаций опубликованы в изданиях, индексируемых в базе данных Web of Science и 20 публикаций в изданиях, индексируемых в базе данных Scopus. Кроме того, четыре статьи из них опубликованы в научных журналах из квартилей Q2 и Q3. Результаты интеллектуальной деятельности отражены в пяти свидетельствах о регистрации программ для ЭВМ. Публикации по теме работы опубликованы в

научных журналах: «Доклады академии наук», «Труды института системного анализа РАН», «Информатика и ее применения», «Искусственный интеллект и поддержка принятия решений», «Информационные технологии и вычислительные системы», «Procedia Engineering» и «Communications in Computer and Information Science». Результаты интеллектуальной деятельности отражены в следующих документах: свидетельство о государственной регистрации программы для ЭВМ Вохминцев А.В., Кобер В.И. Программная система восстановления изображений от мультипликативных интерференций на основе микросканирования камеры № 2015662478; свидетельство о государственной регистрации программы для ЭВМ Вохминцев А.В., Ручай А.Н. Централизованная система дистанционной некооперативной избирательной мультибиометрической аутентификации личности № 2015662479; свидетельство о государственной регистрации программы для ЭВМ Вохминцев А.В., Мельников А.В. Программная система обнаружения объектов на основе дескриптора локальной гистограммы ориентированных градиентов № 2015662752; свидетельство о государственной регистрации программы для ЭВМ Вохминцев А.В., Кобер В.И. Адаптивная система одновременной навигации и составления карты в неизвестном трехмерном пространстве на основе динамического сопоставления изображений и регистрации облаков точек № 2019660919. свидетельство о государственной регистрации программы для ЭВМ Вохминцев А.В., Кобер В.И. Комбинированный метод регистрации мультисенсорных данных на основе точного решения вариационной подзадачи итеративного алгоритма ближайших точек для аффинных и ортогональных преобразований № 2019660920.

Структура и объем работы. Диссертационная работа включает введение, восемь глав, заключение, список терминов, список сокращений и условных обозначений, список литературы (248 наименований), а также два приложения. Диссертация содержит 298 страниц (255 страниц основного текста), 58 иллюстраций, 17 таблиц.

Основные положения, выносимые на защиту:

1. Методология для решения задачи одновременной навигации и построения карты, основанная на комбинировании визуальных и семантических характеристик неизвестной окружающей среды.

2. Алгоритм поиска особых точек на изображениях последовательности кадров на основе дескриптора локальной гистограммы ориентированных градиентов (ГНГ) с рекурсивной процедурой вычисления.

3. Методы решения вариационной задачи точка-точка и точка-плоскость в замкнутой форме для группы аффинных преобразований и для группы ортогональных преобразований с использованием визуально связанных и семантических характеристик.

4. Метод решения задачи навигации и определения положения мобильной платформы, основанный на комбинировании визуальных и семантических характеристик неизвестной окружающей среды и расширенной Калмановской фильтрации (ЕКФ).

5. Быстрый рекурсивный алгоритм для вычисления ДКП дискретного сигнала в скользящем окне с произвольным шагом.

б. Подход к оптимизации графа положений мобильной платформы и обнаружению «замыканий цикла» с использованием метода корзина слов и двумерного дескриптора на основе гистограмм ориентированных градиентов.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность решения проблемы SLAM для контекстуально сложных крупномасштабных сцен; формулируются цель и задачи работы; представлены положения, выносимые на защиту; изложены краткая характеристика и сведения об апробации работы.

Глава 1. В данной главе исследовано современное состояние поставленной в диссертационной работе проблемы, рассмотрены известные методологии решения проблемы: визуальная (Visual SLAM), на основе частичного фильтра (FAST SLAM), на основе расширенного фильтра Калмана (EKF SLAM), на основе графов (Graph-Based SLAM). Предложенная комбинированная методология наследует некоторые свойства визуальной методологии SLAM (Visual SLAM), которая в свою очередь развивает методологию на основе расширенного фильтра Калмана (EKF SLAM). Определены основные проблемы известных методологий применительно к решению задачи SLAM для контекстуально сложных крупномасштабных сцен, представлена информация об основных конкурентах в данной области, проведен патентный поиск по базе ЦИТИС по данной тематике. В данной главе исследованы известные способы решения задачи автоматической трехмерной реконструкции сцен и задачи планирования траектории применительно к решению поставленной в работе практической задачи спасения людей, рассмотрены некоторые аспекты решения задачи некооперативного распознавания личности по лицу на расстоянии. Для решения задачи трехмерной реконструкции во многих приложениях робототехники часто используется итеративный алгоритм ближайших точек (Iterative Closest Point, ICP), либо его модификации.

Глава 2. В работе предложена методология для решения проблемы SLAM, основанная на комбинировании визуальных и семантических характеристик неизвестной окружающей среды (Рисунок 1).

Методология основана на использовании трех типов информации: визуально связанные характеристики окружающей среды (данные о цвете), трехмерные облака точек (данные о глубине сцены), данные о семантических свойствах окружающей среды. В данной методологии входной информацией являются последовательность RGB-D кадров и критерии останова для алгоритмов. Входные данные для компьютерного моделирования и вычислительных экспериментов получены с множества камер, расположенных на мобильной платформе (камеры видимого диапазона, камеры глубины, RGB-D камеры). Выходной информацией являются реконструкция трехмерной сцены и граф положений мобильной платформы, связанный с координатами трехмерной карты в различные моменты времени движения мобильной платформы. Предлагаемая методология в отличие от известных использует мультисенсорную и семантическую информацию на всех этапах решения проблемы SLAM для повышения его точности и состоит из следующей повторяющейся последовательности шагов:

Шаг 1. Получение данных с визуальных камер.

Шаг 2. Процедура предварительной обработки данных (сжатие изображений и восстановление изображений от помех различной природы).

Шаг 3. Снижение размерности двумерных и трехмерных данных в RGB-D кадре с целью уменьшения количества точек в анализируемых данных.

Шаг 4. Извлечение информации об особых точках на изображении с использованием предложенного в работе двумерного дескриптора на основе локальных гистограмм ориентированных градиентов.

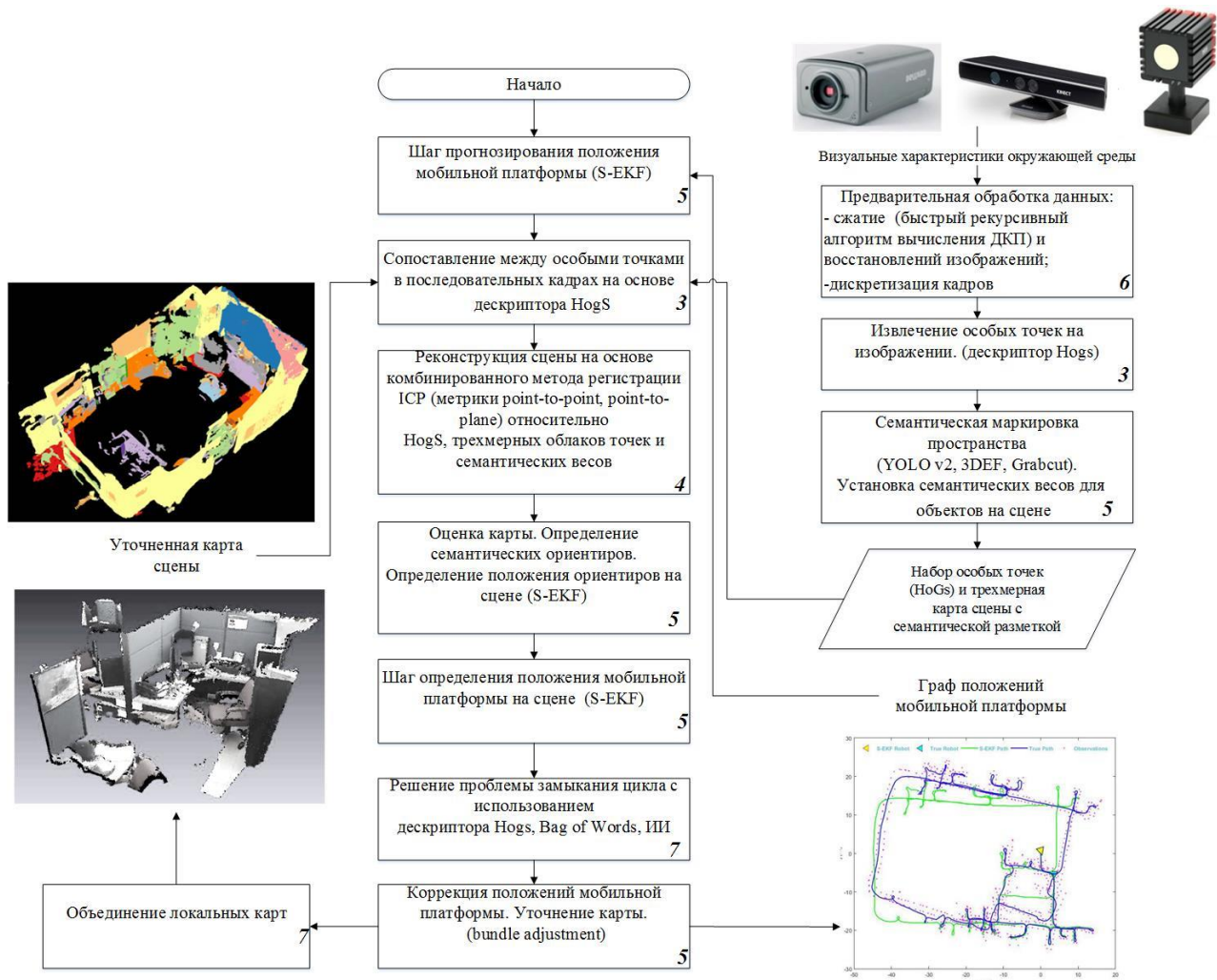


Рисунок 1 – Комбинированная методология SLAM

Шаг 5. Семантическая маркировка 3D пространства на основе 2D изображений и вычисление семантических весов для объектов на сцене.

Шаг 6. Прогнозирование положения мобильной платформы на основе расширенного фильтра Калмана (EKF).

Шаг 7. Сопоставление особых точек на двумерных и трехмерных данных в последовательности кадров.

Шаг 8. Реконструкция сцены на основе комбинированного метода ICP с метрикой точка-точка и метрикой точка-плоскость для группы аффинных и группы ортогональных преобразований соответственно с использованием информации о семантических весах объектов.

Шаг 9. Оценка карты и определение положения навигационных ориентиров на сцене.

Шаг 10. Вычисление положения мобильной платформы на сцене с использованием EKF: вычисление сдвига мобильной платформы на трехмерной сцене и сопоставление данных о трехмерной сцене, полученных на текущем шаге, с данными, которые были получены на предыдущем шаге.

Шаг 11. Решение проблемы «замыкания цикла» с использованием предложенного дескриптора, известного подхода корзина слов и предложенного метода для поиска визуального подобия на основе инвертированного индекса.

Шаг 12. Коррекция положения мобильной платформы на трехмерной сцене и уточнение карты. Оптимизация графа положений платформы.

Шаг 13. Объединение локальных карт. Изменение трехмерной карты окружающей среды, используя результаты, полученные на Шагах 11-12.

Визуально связанные характеристики сопоставляются с использованием предложенного в работе алгоритма сопоставления данных на основе дескриптора ДЛ ГНГ (глава 3). Трехмерные данные сопоставляются с использованием метода ближайших соседей. Информация о семантических характеристиках окружающей среды получена в результате использования известных методов семантической маркировки пространства (YOLO и 3DEF). Ключевым шагом предлагаемой методологии является регистрация данных. Данный шаг связан с комбинированным (совместным) решением вариационной задачи алгоритма ИСР относительно трехмерных облаков точек, визуально связанных и семантических характеристик окружающей среды. Задача регистрации данных заключается в построении модели объекта по множеству его проекций в окружающем пространстве. Данные о характерных признаках изображений были применены на двух основных этапах метода регистрации ИСР: для задания начальных значений и при решении вариационной задачи. Семантические маркеры были связаны со значениями весов слагаемых в рассматриваемом функционале для улучшения сходимости метода регистрации при решении вариационной задачи. В работе предложен точный комбинированный метод решения вариационной задачи для группы аффинных и группы ортогональных преобразований без требования начальной инициализации с использованием информации о семантических характеристиках окружающей среды (глава 4). В работе предложен метод навигации и определения положения мобильной платформы S-EKF SLAM (глава 5), основанный на расширенном фильтре Калмана, использовании мультисенсорной динамической информации, пространственном согласовании последовательных кадров данных и локализации семантических навигационных ориентиров. В главе 6 рассмотрены методы сжатия и восстановления изображений, которые позволяют существенно повысить точность и сходимость всей комбинированной методологии решения проблемы SLAM. В работе предложен подход к решению проблемы «замыкания цикла», а также коррекции движения мобильной платформы на основе известного алгоритма корзина слов (глава 7). Отличительные свойства и преимущества комбинированной методологии представлены в разделе научная новизна.

Глава 3. Точность и сходимость итерационных методов при решении вариационной задачи, связанной с минимизацией функции ошибки, зависит от выбора начальных значений параметров метода ИСР. Как известно, итеративный алгоритм ближайших точек инициализируется начальными значениями параметров геометрического преобразования, такими как поворот, параллельный перенос и изменение масштаба. Выбор параметров инициализации метода ИСР существенно влияет на сходимость данного алгоритма. Для преодоления данного недостатка в работе предлагается использовать визуально связанные характеристики RGB-D кадра, которые позволяют совмещать кадры без требования начальной инициализации. Известные методы сопоставления изображений на основе дескрипторов (SIFT, SURF, ORB) имеют высокую алгоритмическую сложность, поэтому их нельзя использовать для решения поставленной задачи – одновременной навигации и построения карты для крупномасштабных контекстуально сложных сцен. Для обработки визуальных характеристик сцены в данной главе предложен алгоритм сопоставления изображений на основе рекурсивного вычисления гистограмм направленных градиентов (ГНГ) по нескольким круглым скользящим окнам и пирамидальному разложению изображения (ДЛ ГНГ). Пусть $W_i = \{(x, y) \in R^2\}, i = 1, \dots, M$ – набор медианных фильтров (МФ) в виде круглых скользящих окон:

$$W_i = \{(x - x_i)^2 + (y - y_i)^2 \leq r_i\}, \varphi(x, y) = \left[\frac{Q}{360} \arctan\left(\frac{gy}{gx}\right) + \frac{1}{2} \right], \quad (1)$$

где $\{x_i, y_i\}$ – координаты центра медианного фильтра и r_i – радиус i -го диска, соответствующего медианному фильтру, Q – дискретный уровень. В каждом положении i -го МФ из набора вычислим значение градиентов с использованием оператора Собеля. Далее, используя значения модулей градиентов $\{Mag_i(x, y): (x, y) \in W_i\}$ и величины углов направлений $\{\varphi_i(x, y): (x, y) \in W_i\}$, вычислим ГНГ в каждом положении i -го МФ:

$$HOG_i(\alpha) = \begin{cases} \sum_{(x,y) \in W_i} \delta(\alpha - \varphi_i(x, y)), & \text{if } Mag_i(x, y) \geq Med \\ 0, & \text{otherwise} \end{cases}, \quad (2)$$

где $\alpha = \{0, \dots, Q - 1\}$ – порядковый номер дискретного уровня ГНГ, Med – медиана модулей градиентов МФ, $\delta(z) = \{1, \text{если } z = 0; 0 \text{ иначе}\}$ и представляет из себя символ Кронекера. В исследовании используется корреляционный оператор, при помощи которого осуществляется процедура сравнения объектов на анализируемой двумерной сцене с соответствующей ей ГНГ эталонного изображения. Данный оператор позволяет обеспечить инвариантность дескриптора к преобразованию «поворот». Далее введем формулу для определения нормализованной центрированной ГНГ эталонного изображения

$$\overline{HOG_i^R}(\alpha) = \frac{HOG_i^R(\alpha) - Mean^R}{\sqrt{Var^R}}, \quad (3)$$

где $Mean^R$ – среднее значение ГНГ, Var^R – дисперсия ГНГ. Тогда для каждого i -го МФ в позиции k можно определить корреляционную функцию

$$C_i^k(\alpha) = IFT \left[\frac{HS_i^k(\omega)HR_i^*(\omega)}{\sqrt{Q \sum_{q=0}^{Q-1} (HOG_i^k(q))^2 - (HS_i^k(0))^2}} \right], \quad (4)$$

где $HS_i^k(\omega)$ – преобразование Фурье ГНГ внутри i -го МФ входной сцены, $HR_i(\omega)$ – преобразование Фурье $\overline{HOG_i^R}(\alpha)$,* для i -го преобразования Фурье обозначает «комплексное сопряжение». Для определения подобия двух ГНГ применяется корреляционный пик $P_i^k = \max_{\alpha} \{C_i^k(\alpha)\}$. Предложенный двумерный дескриптор обладает лучшими характеристиками по сравнению с известными дескрипторами при малых поворотах вне области сцены.

Глава 4. Во многих приложениях робототехники для построения трехмерных моделей по картам глубины используется алгоритм ИСР. Главный этап алгоритма регистрации ИСР связан с поиском соответствующего геометрического преобразования (ортогонального или аффинного), которое наилучшим образом совмещает два облака точек в разных кадрах для выбранной метрики (вариационная задача алгоритма). Наиболее распространенными вариантами ИСР являются алгоритмы с метриками точка-точка и точка-плоскость. Точность реконструкции трехмерной сцены существенно зависит от выбора метрики для оценки геометрического преобразования и метода решения вариационной задачи. Результат применения итерационных методов для решения задачи минимизации выбранного функционала зависит от правильности выбора начального приближения параметров геометрического преобразования. Использование решений вариационной задачи в замкнутой форме позволяет избежать этих проблем. Выбор класса геометрических преобразований оказывает значительное влияние на результат реконструкции трехмерной сцены. В данной главе предложены методы решения вариационной подзадачи ИСР на основе комбинации данных о визуально связанных характеристиках, трехмерных облаках точек и данных о семантических свойствах окружающей среды.

Рассмотрим вариационную задачу метода регистрации ИСР для аффинных преобразований. Давайте обозначим через $X = \{x_1, \dots, x_n\}$ трехмерные данные об исходном RGB-D кадре и через $Y = \{y_1, \dots, y_m\}$ данные о целевом RGB-D кадре в \mathbb{R}^3 . Пусть отношения между точками в кадрах X и Y такие, что для каждой точки в x_i можно вычислить соответствующую точку в y_i . Тогда итеративный алгоритм ближайших точек (ИСТ) можно рассмотреть как геометрическое преобразование ригидных объектов из X в Y следующего вида

$$Rx_i + T, \quad (5)$$

где R матрица поворота, T вектор параллельного переноса, $i = 1, \dots, n$. Пусть $J(R, T)$ можно представить как функцию следующего вида

$$J(R, T) = \sum_{i=1}^n \| Rx_i + T - y_i \|^2, \quad (6)$$

$$\text{где } R = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}, T = \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix}, x_i = \begin{pmatrix} x_{1i} \\ x_{2i} \\ x_{3i} \end{pmatrix}, y_i = \begin{pmatrix} y_{1i} \\ y_{2i} \\ y_{3i} \end{pmatrix}. \quad (7)$$

Положим, что $\sum_{i=1}^n x_{1i}^2 \neq 0, \sum_{i=1}^n x_{2i}^2 \neq 0, \sum_{i=1}^n x_{3i}^2 \neq 0$, тогда мы можем решить вариационную задачу относительно матрицы поворота R:

$$\frac{\partial J(R)}{\partial r_{1k}} = \sum_{i=1}^n 2 (r_{11}x_{1i} + r_{12}x_{2i} + r_{13}x_{3i} + t_1 - y_{1i})x_{ki} = 0, k = 1,2,3, \quad (8)$$

$$r_{1k} = - \frac{\sum_{i=1}^n (r_{1m}x_{mi} + r_{1n}x_{ni} - y_{1i})x_{ki}}{\sum_{i=1}^n x_{ki}^2}. \quad (9)$$

Введем следующие обозначения для упрощения вида функционала J(R):

$$G_{mi} = x_{mi} - x_{ki} \frac{\sum_{j=1}^n x_{mj}x_{kj}}{\sum_{j=1}^n x_{kj}^2}, \quad (10)$$

$$G_{pi} = x_{ni} - x_{ki} \frac{\sum_{j=1}^n x_{nj}x_{kj}}{\sum_{j=1}^n x_{kj}^2}, G_{ki} = y_{1i} - x_{ki} \frac{\sum_{j=1}^n y_{1j}x_{kj}}{\sum_{j=1}^n x_{kj}^2}.$$

Тогда с учетом обозначений в (10) представим функционал J(R) как

$$J(R) = \sum_{i=1}^n (r_{1m}G_{mi} + r_{1n}G_{pi} - G_{ki})^2 + (r_{21}x_{1i} + r_{22}x_{2i} + r_{23}x_{3i} - y_{2i})^2 + \quad (11)$$

$$+ (r_{31}x_{1i} + r_{32}x_{2i} + r_{33}x_{3i} - y_{3i})^2.$$

Тогда определим частную производную J(R) по r_{1m}

$$r_{1m} = - \frac{r_{1n} \sum_{i=1}^n G_{mi}G_{pi} - \sum_{i=1}^n G_{mi}G_{ki}}{\sum_{i=1}^n G_{mi}^2}. \quad (12)$$

Введем следующие обозначения:

$$q_1 = G_{pi} - \frac{G_{mi} \sum_{j=1}^n G_{mj}G_{pj}}{\sum_{j=1}^n G_{mj}^2}, q_2 = G_{ki} - \frac{G_{mi} \sum_{j=1}^n G_{mj}G_{kj}}{\sum_{j=1}^n G_{mj}^2}. \quad (13)$$

Теперь определим частную производную J(R) по r_{1n} :

$$\frac{\partial J(R)}{\partial r_{1n}} = 2 \sum_{j=1}^n (r_{1n}q_1 - q_2) q_1 = 0, r_{1n} = \frac{\sum_{k=1}^n q_1q_2}{\sum_{k=1}^n q_1^2}. \quad (14)$$

Таким образом определены все элементы первой строки матрицы поворота R: r_{1m}, r_{1k}, r_{1n} . Аналогичным образом получим вторую и третью строку матрицы поворота. Рассмотрим комбинированный метод решения вариационной задачи точка-точка в замкнутой форме для аффинных преобразований

$$J(RV, RD) = \alpha \frac{1}{W} \frac{1}{|A_f|} \sum_{i \in A_f} w_i \|M(RV F_x^i) - M(F_y^i)\|^2 + (1 - \quad (15)$$

$$- \alpha) \frac{1}{W} \frac{1}{|A_d|} \sum_{j \in A_d} w_j \|RDx_j + T - y_j\|^2,$$

где RV, RD – матрицы аффинного преобразования для визуально связанных характеристик сцены и для данных о глубине сцены соответственно; α, W – набор параметров для нормировки данных, подбираемый эмпирическим путем; w_i, w_j – весовые характеристики данных, связанные с семантической маркировкой пространства; A_f – подмножество содержит связи между особыми точками в двух кадрах, A_d – содержит связи между соответствующими точками x_j и y_j в трехмерных облаках точек в двух кадрах данных; F_x^i, F_y^i – визуально связанные характеристики сцены. В общем случае, $RV \neq RD$, в данной работе находится совместное решение вариационной задачи для частного случая, когда $RV = RD = RT$.

Пусть T_{km} – начальная оценка для ИСР с использованием кинематической модели движения камеры, k_{max} , ε – пороги алгоритма ИСР по числу шагов и по величине ошибки E соответственно, RT^* – лучшее преобразование на i -ом шаге метода, δ – порог для точек инлайнеров. Комбинированный метод состоит из шагов:

1. Шаг 1. Вычисление гистограмм направленных градиентов.
2. Шаг 2. Сопоставление между особыми точками F_x^i и F_y^i для выбранных подмножеств. Решение вариационной задачи регистрации данных для визуально связанных характеристик изображения. Получим RT^* и A_f .
3. Шаг 3. Проверка: Если $|A_f| < \delta$, то $RT^* = T_{km}$ и $A_f = \emptyset$. Положим $i=1$.
4. Шаг 4. Определение соответствующих точек A_d в исходном и целевом облаке точек с использованием метода ближайших соседей. Определение весовых коэффициентов для каждой полученной пары A_d .
5. Шаг 5. Решение комбинированной вариационной задачи относительно RT^* , A_f и A_d .
6. Шаг 6. Проверка: $(E(RT^*_i) - E(RT^*_{i+1})) < \varepsilon$ или (Номер итерации $> k_{max}$). Если условие не верно, то возврат на Шаг 4 и $i=i+1$, иначе получено искомое преобразование RT^* .

Для проведения компьютерного моделирования было выбрано четыре набора данных из базы данных NYU Depth Dataset V2, содержащие фрагменты крупномасштабных сцен помещений: Classrooms, Living Rooms (1/4), Offices (1/2), Offices (2/2). В результате серии тестов установлены зависимости скорости сходимости методов от выбора ошибки метрики и условий проведения экспериментов (Рисунок 2). Было установлено, что в контролируемых условиях предложенный метод регистрации данных имеет сходимость близкую к методу, который использует метрику точка-плоскость (Рисунок 2, а) в неконтролируемых условиях (при неравномерном освещении) предложенный метод показывает лучшую сходимость, чем указанные выше методы регистрации данных (Рисунок 2, б). Был проведен сравнительный анализ зависимости скорости сходимости методов регистрации от выбора типа дескриптора и условий проведения экспериментов (Рисунок 3). Было установлено, что в контролируемых условиях тип используемого дескриптора имеет ограниченное влияние на сходимость метода регистрации данных (Рисунок 3, а), в неконтролируемых условиях использование дескриптора ДЛ ГНГ позволяет получить лучшую сходимость в сравнении с другими дескрипторными методами: предложенный метод регистрации сходиться уже после 11 итераций, тогда как при использовании дескриптора ORB метод сходиться только после 16 итераций (Рисунок 3, б). Известно, что метрика точка-плоскость превосходит метрику точка-точка с точки зрения точности и скорости сходимости. Решение вариационной задачи минимизации функционала в замкнутой форме как условной вариационной задачи с ограничением на принадлежность решения классу ортогональных матриц приводит к необходимости работать с существенно нелинейным подмножеством ортогональных матриц многообразия матриц линейных преобразований

$$J(R) = \sum_{i=1}^n ((Rx^i - y^i n^i))^2, \quad (16)$$

где \langle , \rangle есть скалярное произведение, R – матрица аффинного преобразования в однородных координатах, x^i есть точка в облаке точек X , y^i есть точка в облаке точек Y , n^i есть унитарная нормаль к касательной плоскости $T(y^i)$.

$$R = \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{pmatrix}, x^i = \begin{pmatrix} x_1^i \\ x_2^i \\ x_3^i \\ 1 \end{pmatrix}, y^i = \begin{pmatrix} y_1^i \\ y_2^i \\ y_3^i \\ 1 \end{pmatrix}, n^i = \begin{pmatrix} n_1^i \\ n_2^i \\ n_3^i \\ 1 \end{pmatrix}. \quad (17)$$

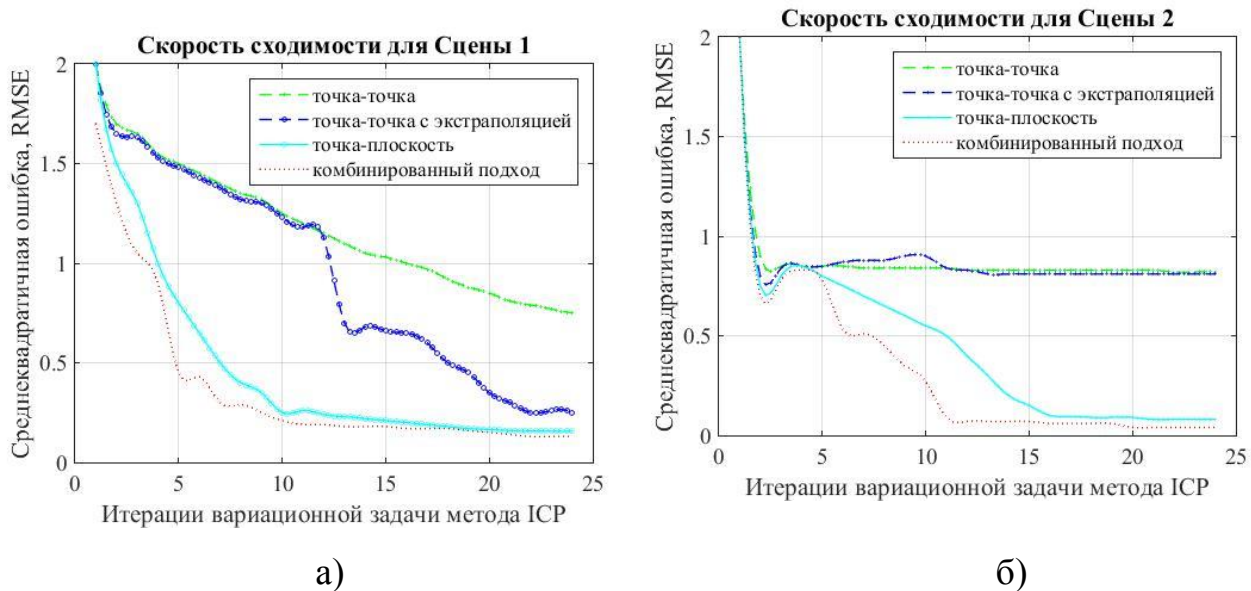


Рисунок 2 – Сравнение скорости сходимости в зависимости от ошибки метрики и условий наблюдения: (а) сходимость в контролируемых условиях, (б) сходимость в неконтролируемых условиях

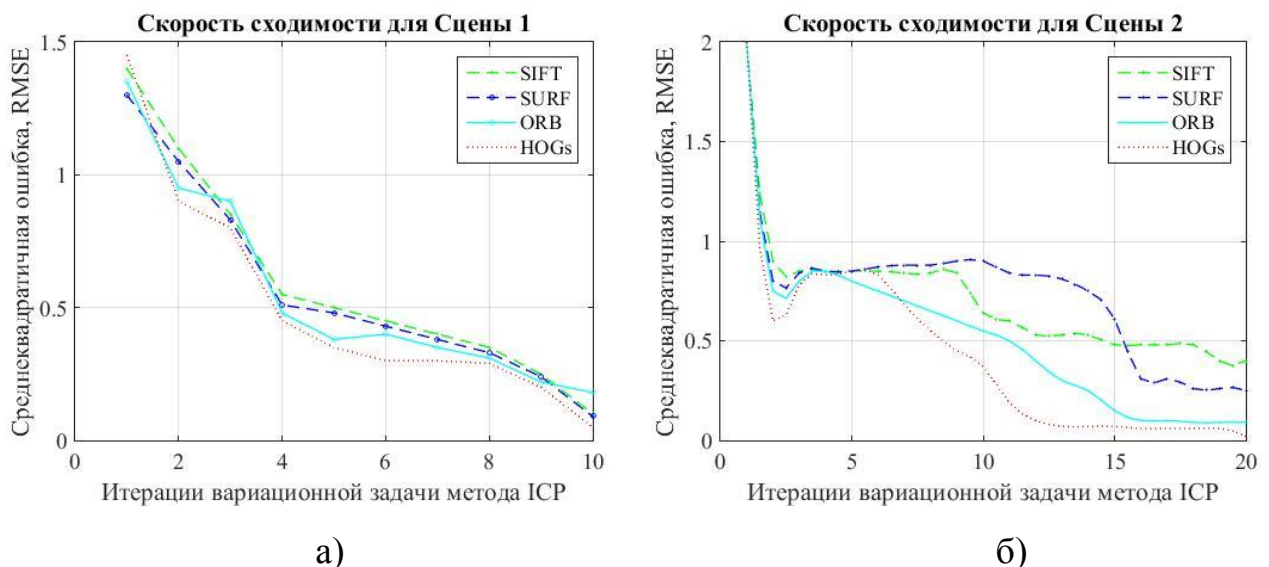


Рисунок 3 – Сравнение скорости сходимости в зависимости от типа дескриптора и условий наблюдения: (а) сходимость в контролируемых условиях, (б) сходимость в неконтролируемых условиях

Рассмотрим функцию $J(R)$

$$\begin{aligned}
J(R) &= \sum_{i=1}^n (\langle Rx^i - y^i, n^i \rangle)^2 = \sum_{i=1}^n (\langle Rx^i, n^i \rangle - \langle y^i, n^i \rangle)^2 = \\
&= \sum_{i=1}^n (\langle Rx^i, n^i \rangle)^2 - 2\langle Rx^i, n^i \rangle \langle y^i, n^i \rangle + (\langle y^i, n^i \rangle)^2.
\end{aligned} \tag{18}$$

Слагаемое $\langle y^i, n^i \rangle$ является постоянным относительно R , следовательно, вариационная задача (18) может быть представлена в следующем виде

$$\sum_{i=1}^n (\langle Rx^i, n^i \rangle)^2 - 2\langle Rx^i, n^i \rangle \langle y^i, n^i \rangle + (\langle y^i, n^i \rangle)^2 \rightarrow \min_R. \tag{19}$$

Рассмотрим компоненты градиента относительно переменной t

$$\frac{\partial J}{\partial t_j} = \sum_{i=1}^n 2\langle Rx^i, n^i \rangle n_j^i - 2n_j^i \langle y^i, n^i \rangle = 0, i, j = 1, \dots, 3. \tag{20}$$

Аналогично найдем частные производные $J(R)$ относительно r_{ij}

$$\frac{\partial J}{\partial r_{ij}} = \sum_{k=1}^n 2\langle Rx^k, n^k \rangle x_j^k n_i^k - 2x_j^k n_i^k \langle y^k, n^k \rangle = 0. \tag{21}$$

Рассмотрим комбинированный метод решения вариационной задачи точка-плоскость в замкнутой форме для ортогональных преобразований с использованием визуально связанных характеристик

$$\begin{aligned}
J(RV, RD) &= \alpha \frac{1}{W|A_f|} \sum_{i=1}^n w_i (\langle RV * F_x^i - F_y^i n^i \rangle)^2 + (1 - \alpha) \frac{1}{W|A_d|} \sum_{i=1}^n w_i (\langle RD * \\
&\quad * x^i - y^i n^i \rangle)^2.
\end{aligned} \tag{22}$$

Обозначим $\eta_1 = \alpha \frac{1}{W|A_f|}$ и $\eta_2 = (1 - \alpha) \frac{1}{W|A_d|}$. Представим функцию $J(RV, RD)$ в следующем виде

$$\begin{aligned}
\eta_1 \sum_{i=1}^n (\langle RV F_x^i, n^i \rangle)^2 - 2\langle RV F_x^i, n^i \rangle \langle F_y^i, n^i \rangle + \eta_2 \sum_{i=1}^n (\langle RD x^i, n^i \rangle)^2 - \\
- 2\langle RD x^i, n^i \rangle \langle y^i, n^i \rangle \rightarrow \min_{RV, RD}.
\end{aligned} \tag{23}$$

Рассмотрим компоненты градиента относительно переменных t и b :

$$\frac{\partial J}{\partial t_j} = \sum_{i=1}^n 2\langle RD x^i, n^i \rangle n_j^i - 2n_j^i \langle y^i, n^i \rangle = 0, j = 1, 2, 3, \tag{24}$$

$$\frac{\partial J}{\partial b_j} = \sum_{i=1}^n 2\langle RV F_x^i, n^i \rangle n_j^i - 2n_j^i \langle F_y^i, n^i \rangle = 0, j = 1, 2, 3. \tag{25}$$

Используем свойство трассировки матрицы, тогда получим:

$$\sum_{i=1}^n Tr \left(RD \cdot (n_j XN)^i \right) = Tr \left(RD \left(\sum_{i=1}^n (n_j XN)^i \right) \right) = \alpha_j, \tag{26}$$

$$\sum_{i=1}^n Tr \left(RV \cdot (n_j FXN)^i \right) = Tr \left(RV \left(\sum_{i=1}^n (n_j FXN)^i \right) \right) = \beta_j. \tag{27}$$

Давайте обозначим через SV^j и SD^j , $j = 1, 2, 3$ следующие матрицы:

$$\sum_{i=1}^n (n_j XN)^i = SD^j, \sum_{i=1}^n (n_j FXN)^i = SV^j. \tag{28}$$

Найдем частные производные функционала $J(RD)$ относительно r_{ij}, f_{ij} :

$$\frac{\partial J}{\partial r_{ij}} = \sum_{k=1}^n 2 \langle RD x^k, n^k \rangle x_j^k n_i^k - 2 x_j^k n_i^k \langle y^k, n^k \rangle = 0, \quad (29)$$

$$\frac{\partial J}{\partial f_{ij}} = \sum_{k=1}^n 2 \langle RV F_x^k, n^k \rangle F_{jx}^k n_i^k - 2 F_{jx}^k n_i^k \langle F_y^k, n^k \rangle = 0. \quad (30)$$

Частная сумма $\sum_{k=1}^n x_j^k n_i^k \langle y^k, n^k \rangle$ есть константа относительно RD , а частная сумма $\sum_{i=1}^n F_{jx}^k n_i^k \langle F_y^k, n^k \rangle$ – относительно RV . Пусть $\sum_{k=1}^n x_j^k n_i^k \langle y^k, n^k \rangle = \alpha_{ij}$ и $\sum_{k=1}^n F_{jx}^k n_i^k \langle F_y^k, n^k \rangle = \beta_{ij}$. Представим решение комбинированной вариационной задачи для ортогональных преобразований в матричной форме:

$$\begin{cases} Tr(RD \cdot SD^j) = \alpha_j \\ Tr(RV \cdot SV^j) = \beta_j \\ \sum_{k=1}^n Tr(RD \cdot SD^{ij}) = \alpha_{ij} \Rightarrow Z \cdot \Omega = \theta, \\ \sum_{k=1}^n Tr(RV \cdot SV^{ij}) = \beta_{ij} \end{cases} \quad (31)$$

где матрица Z имеет размер 18×24 и состоит из элементов SD^{ij} и SV^{ij} ; θ есть матрица из 24 столбцов $\alpha_j, \beta_j, \alpha_{ij}$ и β_{ij} . Комбинированный метод для ортогональных преобразований состоит из шагов:

1. Шаг 1. Вычисление ГНГ на изображениях.

2. Шаг 2. Сопоставление между особыми точками F_x^i и F_y^i для выбранных подмножеств. Решение вариационной задачи регистрации данных для визуально связанных характеристик изображения. Получим RT^* и A_f .

3. Шаг 3. Проверка: Если $|A_f| < \delta$, то $RT^* = T_{km}$ и $A_f = \emptyset$. Положим $i=1$.

4. Шаг 4. Определение соответствующих точек A_d в исходном и целевом облаке точек с использованием метода ближайших соседей. Определение весовых коэффициентов для каждой пары A_d .

5. Шаг 5. Рассмотрим компоненты градиента относительно переменных t и b (векторы переноса) и r_{ij}, f_{ij} (матрицы поворота) соответственно. Получим проекцию элемента многообразия матриц линейных преобразований на подмногообразии ортогональных матриц $F_{jx}^k n_i^k F X N^k, x_j^k n_i^k X N^k$ и $n_j^i X N, n_j^i F X N$.

6. Шаг 6. Решение комбинированной вариационной задачи относительно RT^*, A_f и A_d .

7. Шаг 7. Проверка: $(E(RT^*_i) - E(RT^*_{i+1})) < \varepsilon$ или (Номер итерации $> k_{max}$). Если условие не верно, то возврат на Шаг 4 и $i=i+1$, иначе получено искомое преобразование RT^* .

В результате серии тестов установлены зависимости скорости сходимости методов от выбора ошибки метрики и условий проведения экспериментов (Рисунок 4). Был проведен сравнительный анализ зависимости скорости сходимости

методов регистрации для ортогональных преобразований от выбора типа дескриптора и условий проведения экспериментов (Рисунок 5). Было установлено, что в контролируемых условиях тип используемого дескриптора имеет ограниченное влияние на сходимость метода регистрации данных (Рисунок 5, а), в неконтролируемых условиях использование дескриптора ДЛ ГНГ позволяет получить лучшую сходимость и точность в сравнении с другими дескрипторными методами: предложенный метод регистрации сходиться уже после 10 итераций, тогда как при использовании дескриптора ORB предложенный метод регистрации сходиться только после 14 итераций (Рисунок 5, б). Проведенные вычислительные эксперименты показали значительное повышение точности предложенного метода ICP для разного класса сцен.

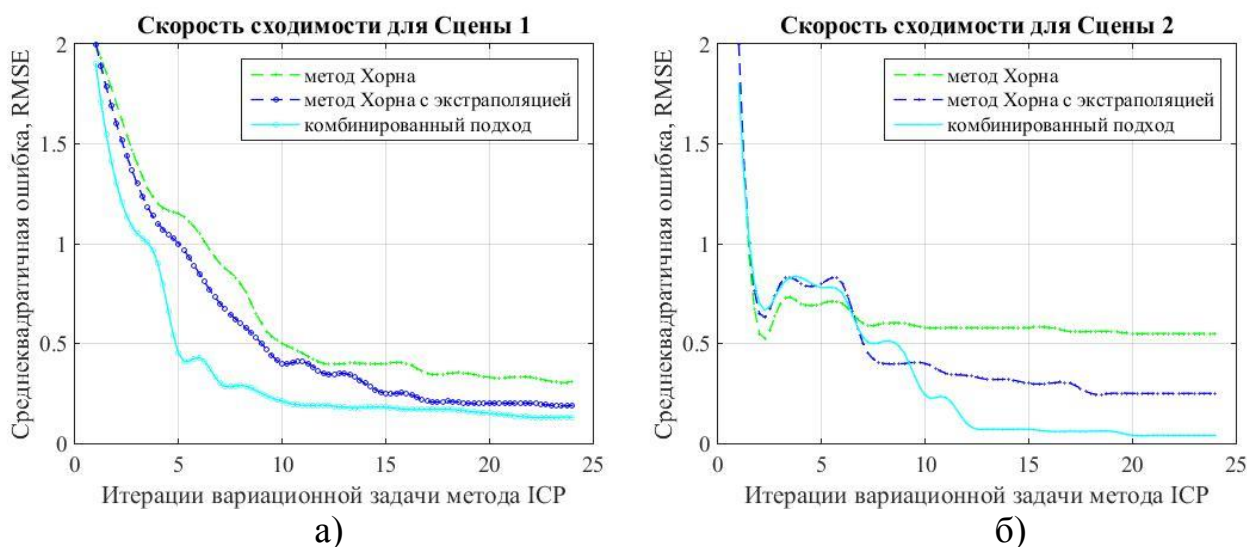


Рисунок 4 – Исследование сходимости в зависимости от ошибки метрики и условий наблюдения: (а) скорость сходимости в контролируемых условиях, (б) скорость сходимости в неконтролируемых условиях

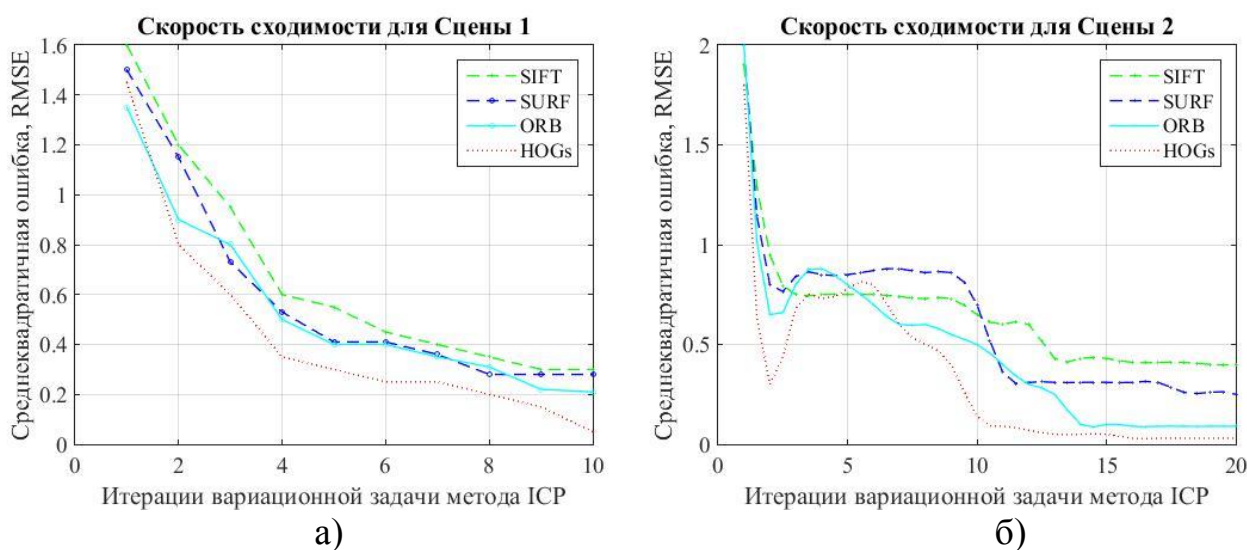


Рисунок 5 – Исследование сходимости в зависимости от выбора типа дескриптора и условий наблюдения: (а) скорость сходимости в контролируемых условиях, (б) скорость сходимости в неконтролируемых условиях

Глава 5. В данной главе предложен метод определения динамического положения мобильной платформы (S-EKF SLAM) на основе расширенной Калмановской фильтрации (EKF), истории положений мобильной платформы, семантических ориентиров и совмещении полученных трехмерных карт. Входной информацией для каждой итерации метода S-EKF SLAM является множество неточных измерений положений платформы на сцене для некоторого дискретного момента времени k , выходной информацией для каждой итерации является последовательность оценок положений мобильной платформы на сцене.

В исследовании используется детектор объектов (YOLO) и семантический подход сегментации, основанный на 3D-запутанных лесах (3-D Entangled Forests 3DEF) для определения принадлежности объектов к классам предметной области (семантической маркировки окружающего пространства). Метод 3DEF позволяет точно сегментировать многие структурные элементы сцены и другие объекты, но плохо классифицирует плоские классы объектов (различные малые формы интерьера). Для преодоления данного недостатка в работе используется детектор объектов YOLO. Для интеграции детектора объектов YOLO и метода 3DEF необходимо: выделить объект в кадре при помощи детектора; внутри ограничительной рамки детектора провести процедуру сегментации объекта вдоль его контуров на основе метода Grabcut. Объединение результатов семантической сегментации метода 3DEF и комбинации методов YOLOv3 + Grabcut осуществляется с использованием слияния на основе Байесовского подхода.

Для каждого шага метода S-EKF SLAM (Рисунок 6) определяется положение мобильной платформы и вычисляется оценка ее следующей позиции. Каждый кадр содержит информацию о местоположении особых точек в пространстве, относительно которых можно вычислить положение мобильной платформы на сцене. Далее определяется значение дельта-функции, которая является разностью полученных оценок положения платформы (реальных и прогнозируемых), затем на основе значения дельта-функции обновляются веса особых точек в RGB-D кадре, производится коррекция положения и траектории движения мобильной платформы на сцене. 3D карта представляет собой вектор состояния большой размерности положений мобильной платформы и навигационных ориентиров: $x = [R \ M]^T$, где R – положение платформы и $M = \{L_1, \dots, L_n\}$ – множество навигационных ориентиров, n – текущее количество отображаемых ориентиров. В EKF SLAM данное отображение моделируется гауссовой переменной с использованием матрицы средних значений и матрицы ковариации вектора состояний и обозначается соответственно как \bar{x} and C :

$$\bar{x} = \begin{bmatrix} \bar{R} \\ \bar{M} \end{bmatrix} = \begin{bmatrix} \bar{L}_1 \\ \vdots \\ \bar{L}_N \end{bmatrix}, C = \begin{bmatrix} C_{R|R} & C_{R|M} \\ C_{M|R} & C_{M|M} \end{bmatrix} = \begin{bmatrix} C_{R|R} & C_{R|L_1} & \cdots & C_{R|L_N} \\ C_{L_1|R} & C_{L_1|L_1} & \cdots & C_{L_1|L_n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{L_N|R} & C_{L_n|L_1} & \cdots & C_{L_n|L_n} \end{bmatrix}. \quad (32)$$

В предложенном методе S-EKF SLAM большая часть трехмерной карты инвариантна относительно движения платформы, поэтому можно представить разреженные матрицы якоби со следующей структурой:

$$F_x = \begin{bmatrix} \frac{dF_R}{dR} & 0 \\ 0 & I \end{bmatrix}, F_y = \begin{bmatrix} \frac{dF_R}{dn} \\ 0 \end{bmatrix}. \quad (33)$$

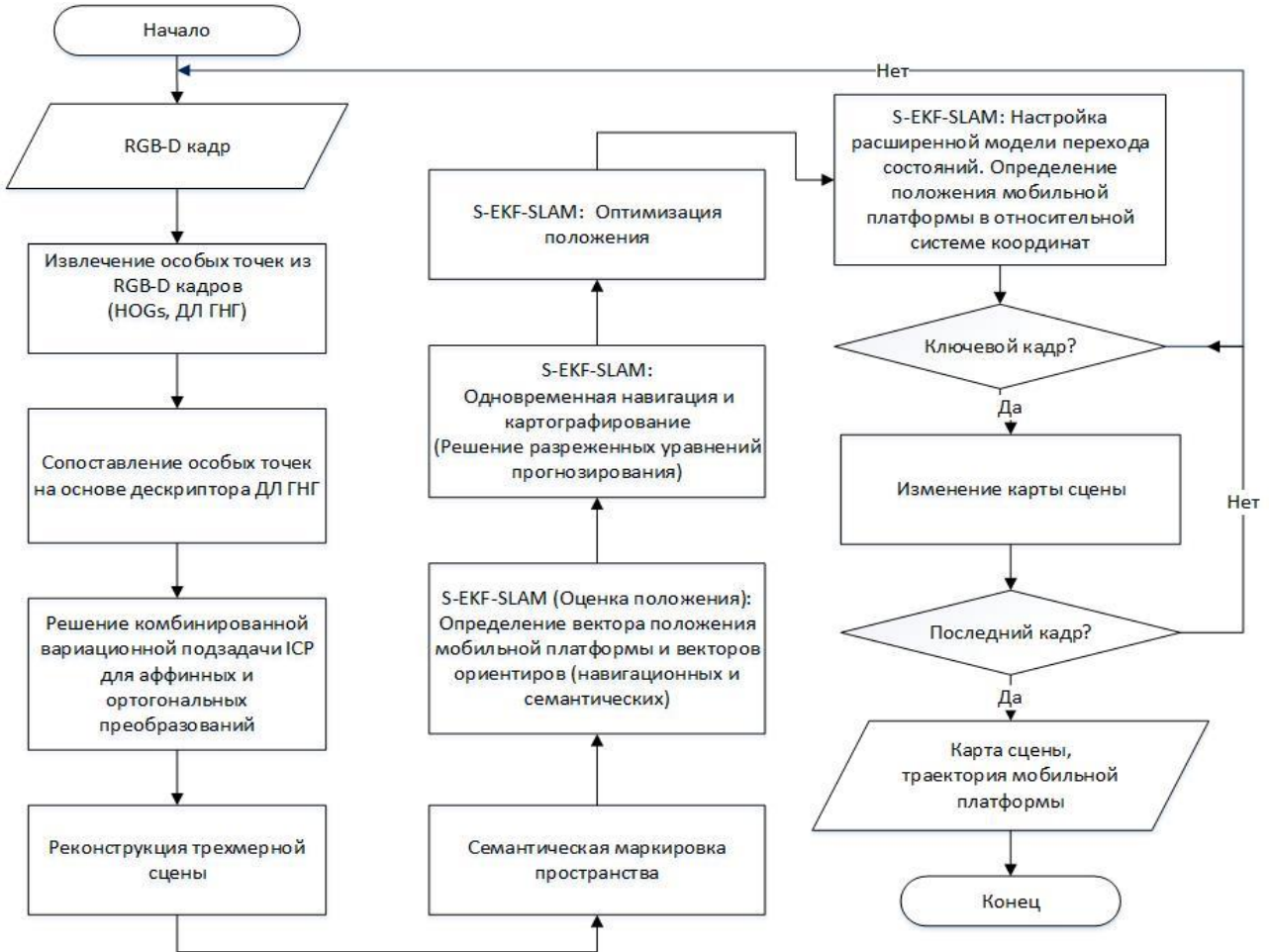


Рисунок 6 – Общая блок-схема метода S-EKF SLAM

Получим разреженные уравнения для прогнозирования траектории движения платформы в пространстве с помощью расширенного фильтра Калмана:

$$\begin{aligned} \bar{R} &\leftarrow f_R(\bar{R}, u, 0), C_{R|R} \leftarrow \frac{dF_R}{dR} C_{R|R} \frac{dF_R^T}{dR} + \frac{dF_R}{dn} N \frac{dF_R^T}{dn}, \\ C_{R|M} &\leftarrow \frac{dF_R}{dR} C_{R|M}, C_{M|R} \leftarrow C_{R|M}. \end{aligned} \quad (34)$$

Состояние исследуемой системы в каждый момент времени состоит из данных о положении и ориентации платформы, а также данных о положении всех семантических ориентиров относительно платформы. Состояние платформы на временном шаге k обозначается $x_v(k)$. Движение платформы моделируется обычным линейным уравнением перехода дискретно-временного состояния или моделью процесса вида $p_i(k+1) = p_i(k) = p_i$. Функцию перехода состояния для i -го семантического ориентира можно представить как $s_i(k+1) = s_i(k) = s_i$. Пусть вектор всех N семантических ориентиров есть $s = [s_1^T \dots s_N^T]^T$, тогда расширенный вектор состояния представим в виде $x(k) = [x_v^T(k) p_1^T \dots p_N^T]^T$. Для получения оценки вектора состояния при движении платформы по серии зашумленных

измерений и семантических ориентиров модель данного процесса представлена в виде матричного уравнения:

$$\begin{bmatrix} x_v(k+1) \\ p_1 \\ \vdots \\ p_N \end{bmatrix} = \begin{bmatrix} JT_v(k+1) & 0 & \cdots & 0 \\ 0 & L_{sp_1} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & L_{sp_N} \end{bmatrix} \text{dist} \left(\begin{bmatrix} x_v(k) \\ p_1 \\ \vdots \\ p_M \end{bmatrix}, \begin{bmatrix} x_v(k) \\ s_1 \\ \vdots \\ s_N \end{bmatrix} \right) + \begin{bmatrix} U_v(k+1) \\ O_{p_1} \\ \vdots \\ O_{p_N} \end{bmatrix} + \begin{bmatrix} V_v(k+1) \\ O_{p_1} \\ \vdots \\ O_{p_N} \end{bmatrix}, \quad (35)$$

где L_{spi} есть $\dim(p_i) \times \dim(p_i)$ – единичная матрица и O_{pi} есть $\dim(p_i)$ – нуль вектор, dist – функция, которая формирует новый вектор навигационных ориентиров p размерностью N из вектора визуальных ориентиров p размерностью M и вектора семантических ориентиров s . Обновление истории перемещений мобильной платформы происходит не на каждом шаге, а если положение платформы существенно изменилось. В EKF SLAM вектор состояния имеет большую размерность, что затрудняет использование методологии для крупномасштабных сцен. Использование семантических ориентиров позволяет сократить размерность вектора состояния при решении проблемы локализации. В работе были проведены компьютерное моделирование для эталонных баз данных и вычислительные эксперименты. Оценка точности предложенного метода S-EKF SLAM для навигации мобильной платформы проводилась в серии экспериментов: первый эксперимент – RGB-D камера переносилась человеком в направлении движения; второй эксперимент – камера Kinect 2.0 и камера MESA SR4000 были установлены на мобильной платформе Husky UGV. Платформы с аналогами систем для тестирования были взяты с сайта openslam.org и были протестированы в различных условиях. На Рисунке 7 представлены результаты второго вычислительного эксперимента.

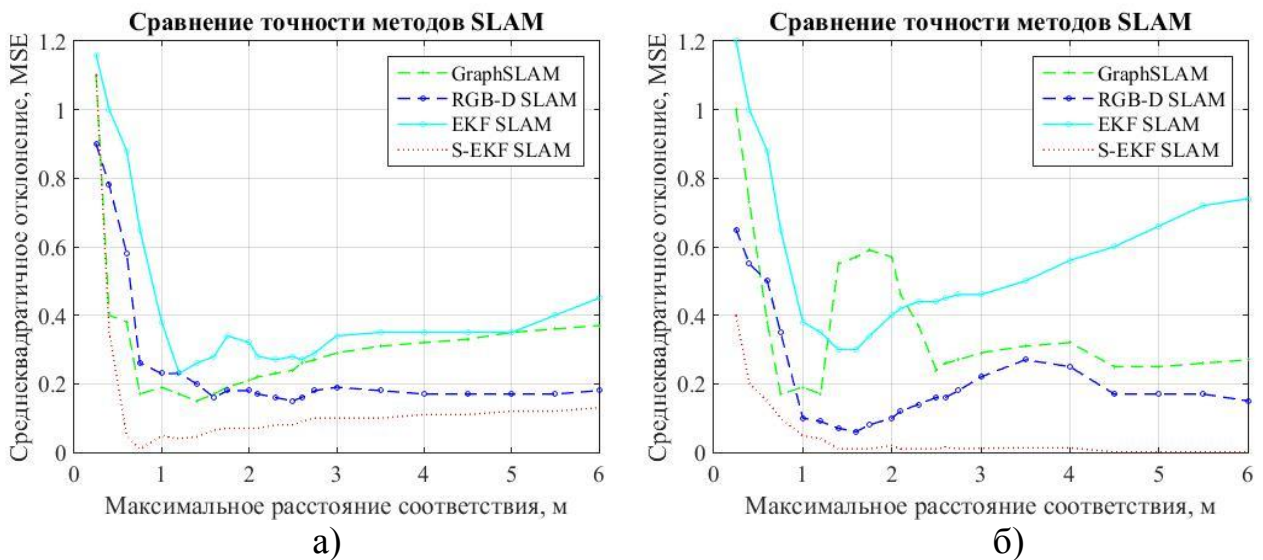


Рисунок 7 – Результаты компьютерного моделирования методов SLAM в контролируемых (а) и в неконтролируемых (б) условиях

Было установлено, что при отсутствии различных шумов и хорошем освещении предложенный метод S-EKF SLAM показывает похожие по точности результаты с классическим Visual SLAM, но при этом предложенный метод имеет лучшую сходимость (Рисунок 7, а). В неконтролируемых условиях (Рисунок 7, б) метод показывает лучшие по точности результаты, чем известные подходы.

Глава 6. В данной главе представлены методы предварительной обработки данных: сжатие изображений и восстановление изображений от помех различной природы. В данной главе предложен быстрый рекурсивный алгоритм для локального вычисления ДКП в окне, скользящем по сигналу с произвольным шагом. Алгоритм обладает меньшей вычислительной сложностью в сравнении с известными рекурсивными и быстрыми алгоритмами ДКП. Текущее ДКП с шагом p может быть определено следующим образом

$$X_s^{kp} = \sum_{n=-N_1}^{N_2} x_{kp+n} \cos\left(\pi \frac{(n + N_1 + 1/2)s}{N}\right), \quad (36)$$

где $\{x_k; k = \dots, -N_1, -N_1+1, \dots, 0, 1, \dots, N_2, N_2+1, \dots\}$ – сигнал бесконечной длины; $\{X_s^{kp}, s = 0, 1, \dots, N-1\}$ – коэффициенты преобразования относительно дискретного времени kp ; $N = N_1 + N_2 + 1$ – длина скользящего окна, где N – произвольное целое число. Коэффициенты ДКП могут быть получены следующим образом: $\{C_0^{kp} = X_0^{kp} / \sqrt{2}; C_s^{kp} = X_s^{kp}, s = 1, \dots, N-1\}$. Рекурсивное уравнение между тремя последовательными скользящими спектрами X_s^{k-2} , X_s^{k-1} и X_s^k имеет вид

$$X_s^k = a_s X_s^{k-1} - X_s^{k-2} + f_s^k, \quad (37)$$

где $f_s^k = (x_{k-N_1-2} - x_{k-N_1-1} + (-1)^s (x_{k+N_2} - x_{k+N_2-1})) \cos(a_s / 2)$, $a_s = 2 \cos(a_s)$ и $a_s = \frac{\pi s}{N}$.

Уравнение (36) – линейное неоднородное разностное уравнение второго порядка. Рекурсивное соотношение между тремя спектрами $X_s^{(k-2)p}$, $X_s^{(k-1)p}$ и X_s^{kp} может быть получено из уравнения (37) следующим образом

$$X_s^{kp} = \begin{cases} 2X_s^{(k-1)p} - X_s^{(k-2)p} + \sum_{m=1}^{p-1} (f_s^{(k-2)p+m+1} + f_s^{kp-m+1}) m + f_s^{(k-1)p+1}, s = 0 \\ 2X_s^{(k-1)p} \cos(a_s p) - X_s^{(k-2)p} + \sum_{m=1}^{p-1} (f_s^{(k-2)p+m+1} + f_s^{kp-m+1}) \frac{\sin(a_s m)}{\sin(a_s)} \\ + f_s^{(k-1)p+1} \frac{\sin(a_s p)}{\sin(a_s)}, s = 1, \dots, N-1 \end{cases}, \quad (38)$$

где $X_s^{k_0}$ и $X_s^{k_1}$ являются начальными спектрами ДКП, вычисляемыми в двух произвольных положениях окна, k_0 и k_1 ($k_0 < k_1$), соответственно. Поскольку окно движется по сигналу с постоянным шагом p , локальные спектры вычисляются в равноудаленных положениях окна kp , где k является целочисленной величиной, а $k > k_1$. Известные рекурсивные алгоритмы ДКП требуют приблизительно $4Np$ арифметических операций для вычисления спектра ДКП в равноотстоящих положениях окна kp . Предложенный алгоритм при $p > 1$ является более эффектив-

ным, чем рекурсивный алгоритм ДКП: дает лучшие результаты, когда шаг скользящего окна p меньше, чем граничное значение, равное 8. В главе представлен непараметрический метод фильтрации и восстановления изображений, искаженных аддитивным фоновым шумом, импульсными помехами и мультипликативными интерференциями, из последовательности пяти наблюдаемых изображений, полученных микросканированием камеры.

Глава 7. Для повышения сходимости предложенной методологии SLAM необходимо исключить из полученного решения (траектории движения) участки с «замыканием цикла». В данной главе предложен подход к оптимизации графа положений мобильной платформы и обнаружению «замыканий цикла» с использованием метода корзина слов и предложенного двумерного дескриптора на основе ГНГ. В данном решении для каждого кадра устанавливается в соответствие предложенный двумерный дескриптор. Для определения «замыкания цикла» на каждом шаге сохраняется история RGB-D кадров и проверяется, совпадает ли текущий кадр с некоторыми из предыдущих кадров. Если текущий кадр является визуально подобным с другим кадром, замеченным в прошлом, то вычисляется преобразование между этими кадрами, и в граф положений добавляется новое ограничение из данного кадра. С помощью новых полученных ограничений положений кумулятивная ошибка графа положений может быть значительно уменьшена. Оптимизация графа положений приводит к глобально выровненному набору кадров. Метод уравнивания связей (bundle adjustment) одновременно оптимизирует граф положений и трехмерную плотную карту. Использование процедуры «замыкания циклов», в общем случае, позволяет существенно улучшить точность локализации мобильной платформы на сцене для длинных циклов (см. Табл. 1 и Табл. 2), для коротких циклов применение процедуры имеет ограниченное влияние, что подтверждается результатами экспериментов.

Таблица 1 – Точность локализации методов SLAM без использования процедуры «замыкания цикла» (для длинных циклов)

Название метода	Среднее отклонение координат, мм	Средне-квадратичное отклонение координат, мм	Среднее отклонение ориентации, град	Средне-квадратичное отклонение ориентации, град
EKF SLAM	215,15	296,98	2,2	1,5
Graph SLAM	629,13	670,33	2,87	1,74
S-EKF SLAM	120,84	198,08	1,71	1,04

Таблица 2 – Точность локализации методов SLAM с использованием процедуры «замыкания цикла» (для длинных циклов)

Название метода	Среднее отклонение координат, мм	Средне-квадратичное отклонение координат, мм	Среднее отклонение ориентации, град	Средне-квадратичное отклонение ориентации, град
EKF SLAM	14,35	7,1	0,3	0,12
Graph SLAM	45,81	30,44	0,65	0,28
S-EKF SLAM	4,05	1,81	0,14	0,09

Предложенный метод S-EKF SLAM на основе комбинирования визуальных и семантических характеристик окружающей среды имеет преимущества с точки зрения точности локализации перед известными методами SLAM на основе EKF. Для повышения производительности подхода в работе используется метод поиска изображений по визуальному подобию с применением инвертированных индексов (ИИ) цветовых гистограмм.

Глава 8. В данной главе рассмотрены практические приложения предложенной методологии для решения проблемы SLAM, рассмотрено применение полученных результатов в системах спасения людей. Система спасения людей состоит из следующих компонент: система одновременной навигации и картографирования; система некооперативного распознавания личности по лицу; система планирования траектории мобильной платформы и позиционирования в неизвестном окружающем пространстве. В диссертационной работе предложены новые методы и подходы: метод дистанционного распознавания личности по лицу с использованием автокодировщика; метод обратной индексации для поиска кандидатов по эталону для распознавания личности по лицу.

ОСНОВНЫЕ ВЫВОДЫ И РЕЗУЛЬТАТЫ

Проведенные в диссертационной работе исследования позволили сформулировать следующие основные выводы и получить практические результаты:

1. Методология для решения задачи одновременной навигации и построения карты, основанная на комбинировании визуальных и семантических характеристик неизвестной окружающей среды, позволяет решать проблему SLAM для контекстуально сложных крупномасштабных сцен в реальном масштабе времени. Предлагаемая методология в отличие от известных использует мультисенсорную визуальную и семантическую информацию на всех этапах решения проблемы SLAM для повышения точности и сходимости получаемого решения. Реализация задач диссертационного исследования выполнена посредством прорывных научных исследований и разработок мирового уровня в области информационных технологий, предложенная комбинированная методология была использована для создания систем навигации и машинного зрения в современных автономных интеллектуальных робототехнических комплексах и системах.

2. Алгоритм представления окрестности особых точек в виде вектора признаков на основе рекурсивного вычисления гистограмм направленных градиентов по нескольким круглым скользящим окнам и пирамидальному разложению изображения позволяет выполнять быстрое сопоставление изображений, инвариантное к повороту и небольшому изменению масштаба для крупномасштабных сцен.

3. Комбинированные методы решения вариационной задачи с метриками точка-точка и точка-плоскость на основе метода ICP для аффинных и ортогональных преобразований позволяют решить важную проблему цифровой обработки информации и робототехники, связанную с реконструкцией точных трехмерных моделей окружающего пространства во время движения автономного интеллектуального объекта управления. Данные о визуально связанных характеристиках были использованы при решении вариационной задачи и позволяют

решить проблему зависимости результата решения вариационной задачи от правильности выбора начальных значений. Предложенный метод регистрации данных для аффинных преобразований создает основу для применения метода Хорна на случай с не ригидными объектами на сцене. Предложены решения вариационной задачи точка-точка для вырожденных случаев, позволяющие проводить точную регистрацию структурных элементов сцены. Некоторые результаты, касающиеся использования информации о характерных точках для решения вариационной задачи точка-плоскость, были известны и ранее, однако явные формулы для совместного решения вариационной задачи в замкнутой форме в поставленной постановке проблемы для решения проблемы SLAM были получены впервые. Комбинированные методы регистрации используются для регистрации облаков точек с произвольным пространственным разрешением и масштабом относительно друг друга. Применение визуально связанных характеристик для решения вариационной задачи точка-точка и точка-плоскость позволяет улучшить точность и сходимость методов регистрации данных в контролируемых и неконтролируемых условиях.

4. Метод решения задачи навигации и определения положения мобильной платформы на сцене, основанный на комбинировании визуальных и семантических характеристик неизвестной окружающей среды и расширенной Калмановской фильтрации (ЕКФ), имеет хорошую сходимость и может использоваться в приложениях, работающих в реальном масштабе времени в контролируемых и неконтролируемых условиях. В работе получены формулы для оценки вектора состояния при движении мобильной платформы по серии зашумленных измерений и семантических навигационных ориентиров. В предлагаемом методе комбинирование информации происходит на уровне выделенных семантических и мультимодальных характеристик, сходимость предложенного метода улучшается за счет использования семантических ориентиров в структуре расширенного фильтра Калмана.

5. Быстрый рекурсивный алгоритм для вычисления ДКП дискретного сигнала в скользящем окне с произвольным шагом основан на рекурсивном соотношении между тремя последовательными локальными спектрами ДКП, которые вычисляются в равноудаленных положениях окна. В работе приведены граничные значения для шага скользящего окна, при которых предлагаемый алгоритм является более эффективным, чем быстрые алгоритмы ДКП. Данный алгоритм используется для локальной пространственно-неоднородной обработки сигналов в спектральной области с произвольным пространственным разрешением.

6. Подход к оптимизации графа положений мобильной платформы и обнаружению «замыканий цикла» на основе метода корзина слов и двумерного дескриптора на основе гистограммы ориентированных градиентов позволяет улучшить сходимость алгоритма обнаружения «замыкания цикла» для длинных циклов, для коротких циклов применение подхода имеет ограниченное влияние.

7. Предложенные методы, алгоритмы и программно-технические решения позволили создать прототип системы одновременной навигации и построения карты лучшего качества, чем известные системы на основе визуальной методологии SLAM и методологии на основе EKF, работающие в аналогичных условиях. Разработанный прототип системы апробирован при решении задачи спасения людей и может быть использован для решения широкого спектра задач: распознавание и слежение за динамическими объектами в заданном пространстве, предсказание траектории их перемещения, согласованное и оптимизированное планирование движений мобильных платформ в заданном пространстве с учетом возможных статических и динамических препятствий.

Статьи в рецензируемых изданиях из списка ВАК

1. Sochenkov, I. Visual duplicates image search for a non-cooperative person recognition at a distance / I. Sochenkov, **A. Vokhmintsev** // Procedia Engineering. – 2015. – Vol.129. – P.440-445.

2. Makovetskii, A. Frequency analysis of gradient descent method and accuracy of iterative image restoration/ A. Makovetskii, **A. Vokhmintsev**, V. Kober, V. Kuznetsov // Communications in Computer and Information Science. – 2015. – Vol. 542. P. 109-118.

3. Kober, V. Recursive calculation of discrete cosine transform for low resolution signal processing / V. Kober, **A. Vokhmintsev**, A. Melnikov // Doklady Mathematics. – 2015. – Vol. 461 (3). – P. 257–260.

4. Соченков И.В., Поиск изображений по визуальному подобию с применением инвертированных индексов цветовых гистограмм / И. В. Соченков, **А. В. Вохминцев**, В.В. Кузнецов, Н. А. Хромов, К.С. Григорьева // Информационные технологии и вычислительные системы. – 2015. – № 4. – С. 86-94.

5. Ручай, А.Н. Разработка централизованной системы избирательной мультибиометрической аутентификации» / А.Н. Ручай, В.В. Кузнецов, А.В. Мельников, **А.В. Вохминцев** // Информационные технологии и вычислительные системы. – 2016. – №.1. – С. 106-116.

6. **Vokhmintsev, A.** A Real-time Algorithm for Mobile Robot Mapping Based on Rotation-invariant Descriptors and ICP / A. Vokhmintsev, K. Yakovlev // Communications in Computer and Information Science. – 2016. – Vol. 661. – P. 338-350.

7. **Vokhmintsev, A.V.** Face recognition based on matching algorithm with recursive calculation of local oriented gradient histogram / A.V. Vokhmintsev, I.V. Sochenkov, V.V. Kuznetsov, D.V. Tikhonkikh // Doklady Mathematics. – 2016. – Vol. 466 (3). – P. 453–459.

8. Соченков, И. Метод обратной индексации для поиска лиц по эталону / И. Соченков, А. Соченкова, **А. Вохминцев** // Искусственный интеллект и принятие решений. 2017. – №.1. – С.50-58.

9. Царегородцев, А.Л. Подход к хранению и обработке первичной слабоструктурированной геолого-геофизической информации / А.Л. Царегородцев, С.С. Слободенюк, С.В. Волошин, **А.В. Вохминцев** // Труды института системного анализа РАН. – 2019. – Т.69 (4). – С. 3-12.

10. **Vokhmintsev, A.V.** Reconstruction of three-dimensional map based on closed form solution of variational problem of multi-sensor data registration / A.V. Vokhmintsev, A.V. Melnikov, K.V. Mironov, V.V. Burlutskiy // *Doklady Mathematics*. – 2019. – Vol. 99 (1). – P. 1–5.

11. **Вохминцев, А.В.** Комбинированные методы навигации и составления карты на основе решения вариационной задачи точка-плоскость ИСР для аффинных преобразований в трехмерном пространстве / А.В. Вохминцев, Мельников А.В., С.В. Пачганов // *Информатика и ее применения*. – 2020. – Т.14 (1). – С. 101-112.

12. **Вохминцев, А.В.** Решение вариационной задачи точка-плоскость ИСР на основе комбинирования визуальных и семантических характеристик трехмерной сцены/ А.В. Вохминцев // *Труды института системного анализа РАН*. – 2020. – Т.70 (1). – С. 3-14.

Публикации, индексируемые в базах данных Web of Science и Scopus

13. **Vokhmintsev, A.** The Knowledge on the Basis of Fact Analysis in Business Intelligence / A. Vokhmintsev, A. Melnikov // *Proceedings of the IFIP TC 5 International Conference, New Prolamat: Digital Product and Process Development Systems (Dresden, Germany October 2013)*. –2013. – P. 354-363.

14. **Vokhmintsev, A.** A fusion algorithm for building three-dimensional maps / A.Vokhmintsev, A. Makovetskii, V. Kober, I. Sochenkov, V. Kuznetsov // *Proceedings of the XXXVIII SPIE's Annual Meeting: Applications of Digital Image Processing (San Diego, USA, August 2015)*. – 2015. – Vol. 9599. –P. 959929-1.

15. Sochenkov, I. A face recognition algorithm based on thermal and visible data / I. Sochenkov, D. Tihonkih, **A. Vokhmintsev**, A. Melnikov, A. Makovetskii // *Proceedings of the XXXIX SPIE's Annual Meeting: Applications of Digital Image Processing (San Diego, USA, September 2016)*. – 2016. – Vol. 9971. – P. 99713F-1.

16. Sochenkov, I. / I. Sochenkov, A. Sochenkova, **A. Vokhmintsev**, A. Makovetskii, A. Melnikov // *Effective Indexing for Face Recognition // Proceedings of the XXXIX SPIE's Annual Meeting: Applications of Digital Image Processing (San Diego, USA, September 2016)*. – 2016. – Vol. 9971. – P. 997124-1.

17. Sochenkova, A. Convolutional neural networks and face recognition task / A. Sochenkova, I. Sochenkov, A. Makovetskii, **A. Vokhmintsev**, A. Melnikov *Proceedings SPIE of the XL Applications of Digital Image Processing (San Diego, USA, August 2017)*. –2017. – Vol. 10396. – P. 103962L-1.

18. **Vokhmintsev, A.** Robot mapping algorithm based on Kalman filtering and symbolic tags / A. Vokhmintsev, T. Botova, I. Sochenkov, A. Sochenkova, A. Makovetskii // *Proceedings SPIE of the XL Applications of Digital Image Processing (San Diego, USA, August 2017)*. – Vol. 10396. – P. 103962I-7.

19. **Vokhmintsev, A.** / A. Vokhmintsev, M. Timchenko, A. Melnikov, A. Kozko, A. Makovetskii // *Robot path planning algorithm based on symbolic tags in dynamic environment // Proceedings of the XL SPIE Applications of Digital Image Processing (San Diego, USA, August 2017)*. – 2017. Vol. 10396. – P. 103962E1-9.

20. **Vokhmintsev, A.** Real-Time visual loop-closure detection using fused iterative close point algorithm and extended Kalman –filter / A. Vokhmintsev, M. Timchenko, A. Karyakina // IEEE Proceedings of the 3rd International Conference on Industrial Engineering, Applications and Manufacturing, ICIEAM, (Saint-Petersburg, Russia, May 2017). – 2017. – P. 1-6.

21. **Vokhmintsev, A.** Simultaneous Localization and Mapping in Unknown Environment using Dynamic Matching of Images and Registration of Point Clouds / A. Vokhmintsev, M. Timchenko, K. Yakovlev // IEEE Proceedings of the 2nd International Conference on Industrial Engineering, Applications and Manufacturing, ICIEAM (Chelyabinsk, Russia May 2016). – 2017. – P. 1-6.

22. Gayanov, R. Transportation of small objects by robotic throwing and catching: applying genetic programming for trajectory estimation / R. Gayanov, K. Mironov, R. Mukhametshin, **A.Vokhmintsev**, D. Kurennov // Proceedings of the 18th IFAC Conference on Technology, Culture and International Stability (Baku, Azerbaijan, September 2018). – 2018. – Vol. 51 (30). – P. 533-537.

23. **Vokhmintsev, A.** Development of methods for selecting features using deep learning techniques based on autoencoders / A. Vokhmintsev, A. Melnikov, M. Timchenko, A. Kozko, A. Makovetskii, A.Kober // Proceedings SPIE of the XLI Applications of Digital Image Processing (San Diego, USA, August-September 2018). – 2018. – Vol. 10752. – P. 1075227-1.

24. **Vokhmintsev, A.** Development of a method for constructing a 3D accurate map of the surrounding environment / A. Vokhmintsev, M. Timchenko, T. Botova, K. Mironov, A. Makovetskii, A. Kober // Proceedings SPIE of the XLI Applications of Digital Image Processing (San Diego, USA, August- September 2018). –2018. – Vol. 10752. – P. 107521X-1.

25. **Vokhmintsev, A.** The New Combined Closed-Solution for 3D Reconstruction of Environment Based on Iterative Closest Point Algorithm / A. Vokhmintsev, A. Melnikov, S. Pachganov, V. Burlutskii // Proceedings of the 7th Scientific Conference on Information Technologies for Intelligent Decision Making Support, ITIDS (Ufa, Russia, May 2019). – 2019, Paris: Atlantis Press. – P.1-6.

Свидетельства об официальной регистрации программы для ЭВМ

26. Свидетельство о государственной регистрации программы для ЭВМ №2015662478 от 02 декабря 2015 г. «Программная система восстановления изображений от мультипликативных интерференций на основе микросканирования камеры» / **Вохминцев А.В.**, Кобер В.И., М.: Роспатент, 2015.

27. Свидетельство о государственной регистрации программы для ЭВМ №2015662752 от 02 декабря 2015 г. «Программная система обнаружения объектов на основе дескриптора локальной гистограммы ориентированных градиентов» / **Вохминцев А.В.**, Мельников А.В., М.: Роспатент, 2015.

28. Свидетельство о государственной регистрации программы для ЭВМ №2015662479 от 02 декабря 2015 г. «Централизованная система дистанционной некооперативной избирательной мультибиометрической аутентификации личности» / **Вохминцев А.В.**, Ручай А.Н., М.: Роспатент, 2015.

29. Свидетельство о государственной регистрации программы для ЭВМ №2019660919 от 15 августа 2019 г. «Адаптивная система одновременной навигации и составления карты в неизвестном трехмерном пространстве на основе динамического сопоставления изображений и регистрации облаков точек» / **Вохминцев А.В.**, Кобер В.И., М.: Роспатент, 2019.

30. Свидетельство о государственной регистрации программы для ЭВМ №2019660920 от 15 августа 2019 г. «Комбинированный метод регистрации мультисенсорных данных на основе точного решения вариационной подзадачи итеративного алгоритма ближайших точек для аффинных и ортогональных преобразований» / **Вохминцев А.В.**, Кобер В.И., М.: Роспатент, 2019.

Прочие публикации

31. **Vokhmintsev, A. V.** Semantic Processing of the Annotated Image Databases / A. V. Vokhmintsev, A. V. Melnikov // Proceedings of the 16th International Workshop on Computer Science and Information Technologies (Sheffield, England, September 2014). – 2014. Vol. 1 – P. 67-71.

32. **Vokhmintsev, A.** A 3D face recognition method based on a modified fast scaling ICP using multisensory biometric data / A. Vokhmintsev // Proceedings of the 17th International Workshop on Computer Science and Information Technologies (Rome, Italy, September 2015). – 2015. Vol. 1 – P. 35-39.

33. **Vokhmintsev, A.V.** The new combined method of the generation of a three-dimensional dense map of environment based on history of camera positions and the robot's movements / A.V. Vokhmintsev, M.S. Timchenko // Proceedings of the 19th International Workshop on Computer Science and Information Technology (Karlsruhe, Germany, October 2017). – 2017. – Vol. 1. – P. 205-211.

34. **Vokhmintsev, A. V.** Solving the point-plane problem for the class of affine transformations and development of a fast iterative algorithm for registration / A. V. Vokhmintsev, A. V. Melnikov, T.S. Botova // Proceedings of the 20th International Workshop on Computer Science and Information Technologies, (Varna, Bulgaria, September 2018). – 2018. – Vol. 1. – P. 73-79.

35. **Вохминцев, А.В.** Алгоритм одновременной навигации и составления карты мобильным роботом на основе итеративного алгоритма ближайших точек и дескриптора, вычисляемого в круглом скользящем окне / А. В. Вохминцев, С. А. Пачганов // Вестник Югорского государственного университета. – 2018. – Т. 3 (50). – С. 7–14. (Включен в перечень ВАК в 2019).

36. **Vokhmintsev, A.** Solving the point-plane problem for the class of affine transformations and development of a fast iterative algorithm for registration of 3d point clouds / A. Vokhmintsev, A. Melnikov, T. Botova // Industry 4.0. – 2018. – Vol.3 (4). – P. 194-199.

Соискатель



А.В. Вохминцев