

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Челябинский государственный университет»

На правах рукописи



ВОХМИНЦЕВ Александр Владиславович

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

Диссертация на соискание ученой степени доктора технических наук

Специальность 05.13.01 – Системный анализ, управление и обработка
информации (в информационных и технических системах)

Научный консультант:

доктор технических наук, профессор

Мельников Андрей Витальевич

Челябинск – 2020

Аннотация

Разработка динамической системы для решения проблемы одновременной навигации и построения карты окружающей среды в реальном масштабе времени является одной из ключевых проблем в современной робототехнике и машинном зрении, так как на ее решении основано создание автономных интеллектуальных робототехнических комплексов и систем. Научная проблема, на решение которой направлено данное исследование – разработка точных методов для реконструкции трехмерной карты доступной окружающей среды для контекстуально сложных крупномасштабных сцен и для определения положения мобильной платформы в неизвестной среде, обоснование оценок их точности и вычислительной сложности. В рамках диссертационного исследования разработана программно-алгоритмическая модель системы для решения проблемы одновременной навигации и построения карты в неизвестном пространстве в реальном масштабе времени.

Содержание

ВВЕДЕНИЕ.....	7
ГЛАВА 1 СОВРЕМЕННОЕ СОСТОЯНИЕ ИССЛЕДОВАНИЙ	23
1.1 Проблема одновременной навигации и построения карты	25
1.2 Автоматическая трехмерная реконструкция сцен	38
1.3 Некооперативное дистанционное распознавание личности по лицу	44
1.4 Результаты патентного поиска.....	48
1.5 Выводы по главе.....	53
ГЛАВА 2 КОМБИНИРОВАННАЯ МЕТОДОЛОГИЯ ДЛЯ РЕШЕНИЯ ПРОБЛЕМЫ ОДНОВРЕМЕННОЙ НАВИГАЦИИ И ПОСТРОЕНИЯ КАРТЫ.....	55
2.1 Постановка задачи. Современное состояние проблемы SLAM	56
2.2 Комбинированная методология при решении задачи регистрации данных .	61
2.3 Сравнительный анализ методов регистрации данных на основе ICP	74
2.4 Методология решения проблемы SLAM на основе комбинирования визуальных и семантических характеристик неизвестной окружающей среды	77
2.5 Выводы по главе.....	86
ГЛАВА 3 МЕТОДЫ ПОИСКА И СОПОСТАВЛЕНИЯ ОСОБЫХ ТОЧЕК НА ИЗОБРАЖЕНИЯХ НА ОСНОВЕ ДВУМЕРНЫХ ДЕСКРИПТОРОВ	90
3.1 Алгоритм извлечения особых точек на основе дескриптора с рекурсивным вычислением гистограмм направленных градиентов	91
3.2 Алгоритм быстрого сопоставления изображений	95
3.3 Компьютерное моделирование	98
3.4 Выводы по главе.....	101
ГЛАВА 4 РЕКОНСТРУКЦИЯ ТРЕХМЕРНЫХ СЦЕН НА ОСНОВЕ ТОЧНЫХ РЕШЕНИЙ ВАРИАЦИОННОЙ ЗАДАЧИ РЕГИСТРАЦИИ МЕТОДА ICP	102

4.1 Точное решение вариационной задачи ICP для аффинных преобразований	106
4.2 Комбинированное решение задачи точка-точка для аффинных преобразований в трехмерном пространстве.....	112
4.3 Компьютерное моделирование	119
4.4 Точное решение вариационной задачи ICP для ортогональных преобразований	124
4.5 Комбинированное решение вариационной задачи ICP для ортогональных преобразований.....	128
4.6 Точное решение комбинированной вариационной задачи ICP для ортогональных преобразований	132
4.7 Компьютерное моделирование	145
4.8 Выводы по главе.....	149
ГЛАВА 5 МЕТОД РЕШЕНИЯ ЗАДАЧИ НАВИГАЦИИ И ОПРЕДЕЛЕНИЯ ПОЛОЖЕНИЯ МОБИЛЬНОЙ ПЛАТФОРМЫ НА ОСНОВЕ КОМБИНИРОВАННОЙ МЕТОДОЛОГИИ SLAM.....	152
5.1 Решение задачи навигации мобильной платформы в неизвестном пространстве на основе расширенного фильтра Калмана и семантических маркеров	154
5.2 Идентификация семантических маркеров на сцене. Метод решения задачи локализации и навигации S-EKF SLAM.....	159
5.3 Компьютерное моделирование	172
5.4 Алгоритм планирования траектории мобильной платформы в неизвестном пространстве на основе метрико-топологического графа и семантических маркеров	179
5.5 Прототип системы на основе комбинированной методологии SLAM.....	186
5.6 Выводы по главе.....	190
ГЛАВА 6 МЕТОДЫ ВОССТАНОВЛЕНИЯ И СЖАТИЯ ИЗОБРАЖЕНИЙ И СИГНАЛОВ	192

6.1 Рекурсивное вычисление дискретного косинусного преобразования для обработки сигналов с пониженным разрешением.....	194
6.2 Рекурсивный алгоритм для вычисления скользящего ДКП с произвольным шагом.....	196
6.3 Точность вариационных методов восстановления сигналов, искаженных линейным оператором	199
6.4 Компьютерное моделирование	203
6.5 Модель базы данных для хранения информации о трехмерной сцене	208
6.6 Выводы по главе.....	210
ГЛАВА 7. РЕШЕНИЕ ПРОБЛЕМЫ «ЗАМЫКАНИЯ ЦИКЛА» И МЕТОДЫ ПОИСКА ИЗОБРАЖЕНИЯ ПО ВИЗУАЛЬНОМУ ПОДОБИЮ	211
7.1 Поиск изображений по визуальному подобию с применением инвертированных индексов цветовых гистограмм	212
7.2 Инвертированный индекс гистограмм для эффективного поиска изображений по визуальному подобию.....	214
7.3 Решение проблемы «замыкания цикла» на основе метода корзина слов и инвертированного индекса.....	217
7.4 Выводы по главе.....	226
ГЛАВА 8 ПРАКТИЧЕСКИЕ ПРИЛОЖЕНИЯ	227
8.1 Применение комбинированной методологии решения проблемы одновременной навигации и картографирования в задаче спасения людей	229
8.2 Метод дистанционного распознавания личности по лицу с использованием автокодировщика.....	235
8.3 Метод обратной индексации для поиска лиц по эталону в задаче спасения людей	241
8.4 Выводы по главе.....	247

ЗАКЛЮЧЕНИЕ	248
Словарь терминов.....	252
Список сокращений и условных обозначений.....	254
Список использованных источников	256
Приложение А. Свидетельства о государственной регистрации программы для ЭВМ	285
Приложение Б. Акты внедрения.....	290

ВВЕДЕНИЕ

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, 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:

1. Разработка быстрых методов для навигации в неизвестном динамическом пространстве.
2. Разработка интеллектуальных методов навигации на основе семантических свойств окружающей среды.
3. Создание методов для навигации на сценах, искаженных помехами окружающей среды различной природы.
4. 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 движения мобильных платформ в 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 проблем 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, Estrada, J. Neira, J.D. Tardós, В.И. Кобера, Е.Н. Соснова, С.Л. Зенкевича и В.С. Лемпицкого [1-7]. 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 05.13.01).
2. Разработать алгоритм поиска и сопоставления особых точек на основе двумерных изображений для обработки информации о цвете в кадре (соответствует п. 4 паспорта специальности 05.13.01).
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 05.13.01).

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 05.13.01).

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 05.13.01).

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 05.13.01).

7. Произвести апробацию предложенной комбинированной визуальной системы SLAM для решения задачи спасения людей (соответствует п. 5 паспорта специальности 05.13.01).

Объектом исследования диссертационной работы является процесс 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 ICP;
- информация о семантических характеристиках окружающей среды используется при решении вариационной задачи точка-плоскость для класса

ортогональных преобразований, при решении вариационной задачи 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 ICP 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 ICP: 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 ICP 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 ICP для ортогональных преобразований позволяет получить в неконтролируемых условиях лучшую сходимость и точность в сравнении с другими дескрипторными методами: предложенный метод регистрации сходится уже после 10 итерации, тогда как при использовании дескриптора ORB предложенный метод регистрации сходится 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 сходимостъ;
- 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 сходимостъ 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 корзины 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 в работе методов и алгоритмов 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 исследований 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 научно-исследовательской 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». Магистерская программа создана в рамках выполнения грантов для преподавателей магистратур Благотворительного фонда В. Потанина в 2017-2018 гг. и в 2019-2020 гг. (руководитель Вохминцев А.В.).

Апробация 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:

1. 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.).

2. РНФ № 15-19-10010 «Разработка алгоритмической модели технической системы для идентификации личности по мультисенсорным биометрическим данным», 2015-2017 гг. (основной исполнитель, руководитель Кобер В.И.).

3. 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.).

4. 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.).

5. 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 HOGMan, ORB-SLAM, TORO-SLAM, RGBDSlam, EKFMonoSLAM, используя мобильные платформы Odyssey 6 Robotics и Husky UGV.

Публикации. По теме диссертационной работы всего опубликована 31 научная публикация [8-38]: 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 [8-10] и Q3 [13]. 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 принятий решений», «Информационные технологии и вычислительные системы», «Procedia Engineering» и «Communications in Computer and Information Science».

Результаты интеллектуальной деятельности отражены в следующих документах [39-43] (см. Приложение А):

1. 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.

2. Свидетельство о государственной регистрации программы для ЭВМ Вохминцев А.В., Ручай А.Н. Централизованная система дистанционной некооперативной избирательной мультибиометрической аутентификации личности № 2015662479.

3. 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.

4. 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.

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

ГЛАВА 1 СОВРЕМЕННОЕ СОСТОЯНИЕ ИССЛЕДОВАНИЙ

В данной главе исследовано современное состояние поставленной в работе проблемы и рассмотрены известные методологии SLAM. Определены основные проблемы известных методологий применительно к решению проблемы SLAM для контекстуально сложных крупномасштабных сцен, приведено сравнение ожидаемых результатов с мировым уровнем, представлена информация об основных конкурентах в данной области, проведен патентный поиск по базе ЦИТИС по данной тематике.

В работе предложена новая методология SLAM для навигации мобильной платформы в замкнутом неизвестном пространстве. Решение проблемы основано на динамическом сопоставлении и регистрации последовательности RGB-D кадров и вариантах Калмановской фильтрации. Для построения плотных трехмерных карт используется комплекс сенсорных датчиков (камеры визуального диапазона и времяпролетные камеры), которые позволяют получать данные о цвете и глубине на сцене. Предлагаемая трехмерная картографическая система основана на трех компонентах: пространственное совмещение последовательных кадров данных; обнаружение «замыканий цикла»; сопоставление и регистрация последовательностей изображений и разреженных облаков трехмерных точек. На основе истории положений камеры и перемещений мобильной платформы, семантических характеристик сцены, совмещения полученных трехмерных карт глубины, учитывающих точность их наложения и геометрические связи между различными снимками одной сцены, генерируется трехмерная объединенная карта доступной окружающей среды и определяется позиция платформы в относительной системе координат. Для построения траектории движения мобильной платформы из начальной точки в конечную (целевую) точку используются эвристические методы планирования траектории.

Диссертационное исследование направлено на реализацию приоритетных направлений развития Российской Федерации в части перехода к

роботизированным системам, создания систем машинного обучения и искусственного интеллекта, в том числе в сфере обеспечения общественной безопасности и противодействия терроризму. Перспективным направлением применения результатов научно-исследовательских работ является антропоморфная робототехника и связанные с ней области исследований. Система трехмерного машинного зрения, предложенная в исследовании, может быть использована в российских аналогах следующих проектов: LS3 (Boston Dynamics, SoftBank) [44], Rollin' Justin (Institute of Robotics and Mechatronics) [45], Walking robots (Robotics) [46], Baxter и Sawyer (Rethink Robotics) [47], Nextage [48].

На основе предложенных методов и алгоритмов создана система для решения проблемы одновременной навигации и картографирования группой мобильных платформ в неизвестном пространстве в реальном масштабе времени с использованием персональных компьютеров с графическими процессорами и технологий распараллеливания и программирования.

Система одновременной навигации и построения карты на основе комбинированной методологии выполняет следующие функции:

- осуществление навигации в неизвестном окружении;
- построение карты окружающего пространства;
- обнаружение и идентификация субъектов по лицу.

Предлагаемая в диссертационном исследовании методология ориентирована на перспективные рынки робототехнических систем и комплексов, промышленное производство коллоборативных роботов, систем биометрической идентификации личности. В работе рассмотрено применение разработанной комбинированной системы одновременной навигации и построения карты при создании систем спасения людей.

1.1 Проблема одновременной навигации и построения карты

Решение проблемы одновременной навигации и картографирования (Simultaneous localization and mapping, SLAM) [49, 50] связано с одновременным решением двух задач во время движения мобильной платформы: реконструкция карты неизвестной окружающей среды вокруг мобильной платформы и определение положения мобильной платформы в каждый момент времени на карте. Входной информацией для решения данной задачи являются данные, получаемые с различных мультисенсорных датчиков, например, RGB-D камер. В определенные дискретные моменты времени датчики сканируют окружающую среду и отправляют полученные данные в виде кадров для последующей обработки и анализа. Таким образом, кадр является системой отсчета времени в системе SLAM. Выходной информацией для данной задачи являются реконструкция сцены и траектория движения мобильной платформы на сцене (оценка SLAM). Данная задача в общем виде может быть представлена в виде итерационной процедуры, состоящей из следующей последовательности шагов:

Шаг 1. Исследование окружающей среды (трехмерной сцены). Получение информации (визуальной, геометрической, одометрической, семантической) о характеристиках окружающей среды.

Шаг 2. Вычисление сдвига мобильной платформы на трехмерной сцене. Сопоставление данных о трехмерной сцене, полученных на текущем шаге, с данными, которые были получены на предыдущем шаге.

Шаг 3. Определение характерных признаков в текущих данных (например, выделение особых точек на изображении в кадре).

Шаг 4. Процедура сопоставления характерных признаков для текущих данных с характерными признаками для заданной трехмерной сцены за всю историю наблюдений.

Шаг 5. Изменение положения мобильной платформы на основе результатов, полученных на Шаге 4.

Шаг 6. Решение проблемы «замыкания цикла» (проверка полученной траектории движения мобильной платформы на петли).

Шаг 7. Глобальная оптимизация данных. Изменение трехмерной карты окружающей среды, используя результаты, полученные на Шаге 4 и Шаге 5. (данные о положении характерных признаков и мобильной платформы на трехмерной сцене за всю историю наблюдений).

На каждом шаге решения задачи SLAM требуется иметь следующую информацию:

- положение мобильной платформы на сцене в текущий момент времени;
- информацию о навигационных ориентирах на сцене;
- доступную часть карты сцены, которая была получена с начального момента времени, соответствующего исходному положению мобильной платформы на сцене, до текущего момента времени.

В современной робототехнике принято выделять следующие методологии SLAM:

- методология быстрого SLAM на основе частичного фильтра (Particle Filter SLAM). Также в литературе используется альтернативное название FAST SLAM [52];
- методология SLAM на основе расширенного фильтра Калмана (Extended Kalman Filter, EKF) [51, 53];
- визуальная методология SLAM (Visual SLAM) [54-56];
- методология SLAM на основе графов (Graph-Based SLAM) [57].

На рисунке 1.1 представлена укрупненная схема для решения проблемы одновременной навигации и построения карты на основе визуальной методологии. Входной информацией в данном случае является информация от различных мультисенсорных датчиков, например, RGB-D камер, выходной информацией является оценка SLAM, которая представляет реконструкцию двумерной/трехмерной карты и траекторию движения мобильной платформы на

сцене. Каждый шаг проблемы SLAM можно условно разделить на две части: цифровая обработка данных, полученных с мультисенсорных датчиков; оценка состояния и корректировка карты и положения мобильной платформы.



Рисунок 1.1 – Проблема одновременной навигации и построения карты

Существующие проекты в области SLAM размещены на портале <http://openslam.org/>, который обеспечивает исследователей программными платформами на основе перечисленных ранее методологий, среди которых можно выделить следующие успешные коммерческие реализации: HOG-Man, ORB-SLAM, TORO-SLAM, RGBD-Slam, EKF-Mono-SLAM.

Методология SLAM на основе расширенного фильтра Калмана во многом аналогична классическому способу применения метода Extended Filter Kalman для решения задачи локализации мобильной платформы на трехмерной сцене. Однако методологию SLAM на основе EKF невозможно использовать для навигации мобильной платформы на крупномасштабных сценах в реальном масштабе времени. Это обусловлено тем, что, во-первых, при решении задачи SLAM размерность вектора состояния в фильтре Калмана значительно больше, чем при решении задачи локализации. Более того, размерность вектора состояния увеличивается с каждым новым шагом метода SLAM за счет получения информации о новых объектах на сцене и информации о навигационных ориентирах. Во-вторых, в связи с динамическим ростом вектора состояния в

расширенном фильтре Калмана методология решения проблемы SLAM на основе EKF имеет высокую вычислительную сложность и не может быть использована в реальном масштабе времени для крупномасштабных сцен. В-третьих, шум матрицы ковариации в методологии на основе EKF SLAM имеет квадратичную возрастающую зависимость от размера вектора состояния, которая может быть определена следующим образом $(3+2n)*(3+2n)$.

Как было сказано ранее, исторически первыми появились быстрые методы SLAM, которые основаны на применении Байесовской сети и фильтров частиц. Фильтр частиц организован как набор дискретных частиц в пространстве состояний, представляющий распределение вероятности некоторой величины. Далее на каждом следующем шаге алгоритма изменения состояния системы происходит генерация нового распределения вероятности для заданной модели управления мобильным объектом (мобильной платформой). В данном методе условная вероятность определяется на основе теоремы Рао-Блэквелла-Колмогорова. В FAST-SLAM каждая оценка навигационного ориентира задана в виде расширенного фильтра Калмана размером 2×2 . Методология быстрого SLAM имеет ограниченный набор приложений, в которых анализируются и обрабатываются сцены ограниченной размерности, для реконструкции крупномасштабных сцен и навигации на них данная методология не может быть применена. Основные недостатки методологии быстрого SLAM связаны с плохой масштабируемостью фильтра частиц и невозможностью выполнять точное сопоставление частиц в условиях неопределенности.

Большое распространение получили методы на основе методологии Graph-Based SLAM [58]. В данных методах узел разреженного графа соответствует положению объекта (мобильной платформы) и другим элементам карты, а связи в графе соответствуют относительным позициям внутри вектора, задающего положение платформы в пространстве, и относительным местоположениям между положением платформы на сцене и элементами двумерной или трехмерной карты, которые можно наблюдать из данных местоположений. Положение мобильной платформы в пространстве задается входом одометрии [59, 60].

Современные методологии SLAM для крупномасштабных сцен основаны на визуальной методологии SLAM или ее комбинации с методологией на основе EKF SLAM [54-56, 58]. Методологию SLAM в данном случае можно описать в виде следующей схемы:

- локальное построение карты с использованием метрик для оценки траектории и создания локальных карт;
- распознавание положений на карте для решения проблемы «замыкания цикла», либо для решения задачи повторной локализации;
- оптимизация построенной карты и глобальная оптимизация.

На сегодняшний день в области цифровой обработки информации были предложены различные методы и алгоритмы для решения проблемы реконструкции трехмерных сцен на основе различных мультисенсорных датчиков: времяпролетные камеры, камеры визуального диапазона, RGB-D датчики [61-65], 2D и 3D стереоскопические фотоаппараты [66], призматический монокуляр [67]. Известны подходы к решению задачи реконструкции трехмерной сцены с использованием эвристических алгоритмов анализа и обработки цифровых неотсортированных массивов изображений [68-70]. Современные методы трехмерной реконструкции основаны на решении задачи двумерной или трехмерной регистрации данных, при этом различают итеративные методы регистрации данных и не итеративные методы регистрации данных. Во многих приложениях, в том числе при решении задач робототехники, часто используется итеративный алгоритм ближайших точек (Iterative Closest Point, ICP), либо его многочисленные модификации [71-73]. Входной информацией для данного алгоритма является множество данных в виде двумерного или трехмерного облака точек. Выходными данными является геометрическое преобразование, которое позволяет минимизировать расстояние между двумя геометрическим моделями, полученными с разных точек обзора сцены, на основе некоторой метрики. Итеративный алгоритм ближайших точек может завершиться либо по достижении определенного порога точности, задаваемого пользователем, либо по одному или нескольким критериям остановки работы алгоритма, например, максимально

допустимое число итераций. Различают несколько вариантов минимизации функции ошибки:

- многократное создание набора соответствующих точек с использованием текущего преобразования и поиск нового преобразования;
- итеративная минимизация в сочетании с экстраполяцией в пространстве преобразования;
- итеративная минимизация с использованием различных произвольно выбранных подмножеств точек и выбор оптимального преобразования с использованием метрики наименьшей медианы квадратов;
- стохастический поиск наилучшего преобразования с использованием метода имитации отжига.

Выбор метрики и способа минимизации функции ошибки определяет точность и сходимость конкретного варианта метода ICP. Известно, что сходимость и точность классических алгоритмов ICP, предложенных в работах Бесла П., Чена Ю., Хорна Б., может быть существенно улучшена. Детальный обзор различных методов на основе ICP приведен в известных работах Русинкевича Ш. и Левого М. [74, 75], которые посвящены исследованию эффективности методов ICP. Рассмотренные в данной работе методы применительно к задаче одновременной навигации и картографирования (SLAM) на основе визуальной методологии имеют хорошую сходимость и точность на представленных эталонных базах данных [61, 75]. Однако, представленные в работе эталонные базы данных основаны только на данных, полученных с датчиков глубины, тогда как современные методы и системы SLAM используют для обработки и анализа комплекс гетерогенной информации: данные с камер видимого диапазона; данные с камер глубины; данные с камер инфракрасного диапазона; одометрические данные; данные о семантических характеристиках исследуемой трехмерной сцены. Методы, которые исследованы в работах [74, 75], накладывают существенные ограничения на анализируемые данные: сдвиг между последовательностью кадров данных должен быть незначительным; все объекты на трехмерной сцене должны находиться в области обзора различных мультисенсорных датчиков, должны быть доступны

проекции трехмерного объекта с разных точек обзора. Если данные условия не соблюдаются, то предложенные варианты итеративного алгоритма ближайших точек показывают недостаточно высокие результаты по точности трехмерной реконструкции (двумерные или трехмерные облака точек плохо совмещены), либо предложенные варианты не сходятся в принципе. Поэтому, задача создания новых методов реконструкции трехмерных данных на основе разнородной информации становится актуальной для развития робототехники, методов машинного зрения и методов автономной навигации мобильных платформ в пространстве. В данной диссертационной работе предлагается создать новый итеративный метод ближайших точек, который будет основан на комбинировании трех источников информации: данные о глубине сцены, данные о цвете сцены и семантические характеристики сцены.

В результате решения проблемы SLAM определяется траектория движения мобильной платформы на сцене, привязанная к реконструируемой сцене. Для повышения сходимости предложенного метода SLAM необходимо исключить из полученного решения (траектория движения) участки «с замыканием цикла». Современные решения данной проблемы основаны на использовании методов быстрого сопоставления данных, которые пришли в робототехнику и машинное зрение из области обработки текстов на естественном языке (Natural Language Processing) [61, 62]. Наиболее известные методы решения этой проблемы основаны на использовании корзины слов (Bag of Words) [65, 67]. Траектория движения мобильной платформы представляет собой некоторый связанный граф положений. При обнаружении на траектории движения «замыкания цикла», так называемой петли, создается новое ограничение в графе положений, которое определяет новое отношение между последовательностью RGB-D кадров для данной сцены. Для решения задачи одновременной оптимизации графа положений и карты сцены используется метод слепой подстройки [76], кроме классического варианта данного метода известны его многочисленные аналоги и модификации [77-79].

Решение задачи планирования траектории движения мобильной платформы разделяют на два этапа: поиск оптимальной траектории на карте и навигация

платформы вдоль полученной траектории. Рассмотрим существующие методы поиска локально-оптимальной траектории. Для представления карт окружающего пространства используются сетчатые карты занятости и топологические карты. «На сегодняшний день разработаны специальные алгоритмы поиска кратчайшего пути: алгоритм A^* и его модификации: D^* (динамический A^*) [80], FD^* (фокусированный динамический A^*) [81], D^* Lite [82]. Наибольшее распространение в динамическом программировании для определения эффективного пути получили алгоритмы, основанные на следующих графовых моделях: графы видимости [83-85], вероятностные схемы местности [86], графовые модели, основанные на регулярной декомпозиции (ГРД) [87], метрический топологический граф [88]» [180].

В данном исследовании будет предложена новая комбинированная методология на основе визуальной методологии SLAM (Visual SLAM), которая в свою очередь развивает методологию на основе расширенного фильтра Калмана (EKF SLAM). Давайте рассмотрим более подробно архитектуру методологии EKF SLAM (Рис. 1.2) и методологии Visual SLAM (Рис. 1.3).

В методологии EKF SLAM входной информацией являются данные с различных лидаров, например, с лазерных дальномеров и данные с одометрических датчиков. Выходной информацией являются граф положений мобильной платформы и карта сцены, содержащая положения мобильной платформы и навигационных ориентиров. В методологии EKF SLAM проблема SLAM может быть представлена в виде итерационной процедуры, состоящей из следующей последовательности шагов:

Шаг 1. Получение данных выхода одометрии и данных с лидаров.

Шаг 2. Прогнозирование положения мобильной платформы.

Шаг 3. Оценка карты и определение положения навигационных ориентиров на сцене.

Шаг 4. Вычисление положения мобильной платформы на сцене: вычисление сдвига мобильной платформы сцене и сопоставление данных о сцене, полученных на текущем шаге, с данными, которые были получены на предыдущем шаге.

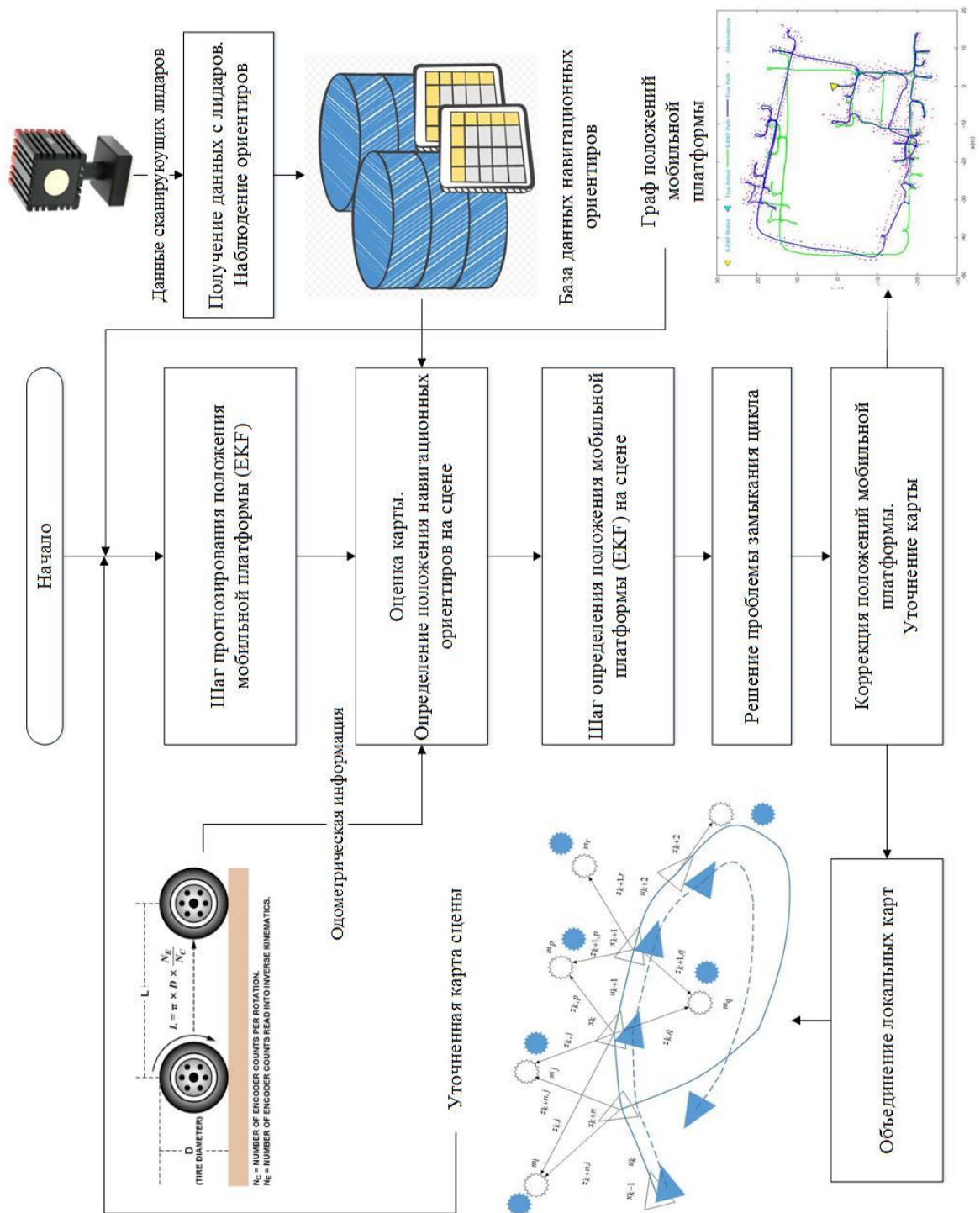


Рисунок 1.2 – Методология SLAM на основе расширенного фильтра Калмана (EKF SLAM)

Шаг 5. Решение проблемы «замыкания цикла».

Шаг 6. Коррекция положения мобильной платформы на сцене и уточнение карты.

Шаг 7. Объединение локальных карт. Изменение карты окружающей среды, используя результаты, полученные на Шаге 5-6.

В методологии Visual SLAM входной информацией являются данные с одометрических датчиков; визуальные данные в виде данных о цвете (изображения) и данных о глубине сцены (трехмерные облака точек, карты глубины), получаемые с различных мультисенсорных датчиков (камеры видимого диапазона, камеры глубины, RGB-D камеры). Выходной информацией являются граф положений мобильной платформы и реконструкция двумерной или трехмерной карты сцены, содержащая особые точки. Между графом положений и картой сцены устанавливается соответствие для вычисления положения мобильной платформы и навигационных ориентиров на сцене в каждый момент времени. В данной методологии проблема SLAM представляет из себя итерационную процедуру, состоящую из следующей последовательности шагов:

Шаг 1. Получение данных выхода одометрии и данных с визуальных камер, например, с RGB-D камеры Kinect.

Шаг 2. Предварительная обработка визуальных данных: восстановление и сжатие изображений, дискретизация 2D и 3D данных. Использование необработанных данных с помехами различной природы при решении задачи сопоставления данных и задачи регистрации данных часто приводит к тому, что результат реконструкции сцен имеет неудовлетворительное качество.

Шаг 3. Извлечение особых точек на изображении с использованием одного из известных дескрипторов (SIFT, SURF, ORB).

Шаг 4. Прогнозирование положения мобильной платформы на основе расширенного фильтра Калмана (EKF).

Шаг 5. Установление соответствия между особыми точками (сопоставление данных) в последовательности кадров.

Шаг 6. Реконструкция сцены на основе известных методов регистрации, например, на основе итеративного алгоритма ближайших точек (ICP).

Шаг 7. Оценка карты и определение положения навигационных ориентиров на сцене.

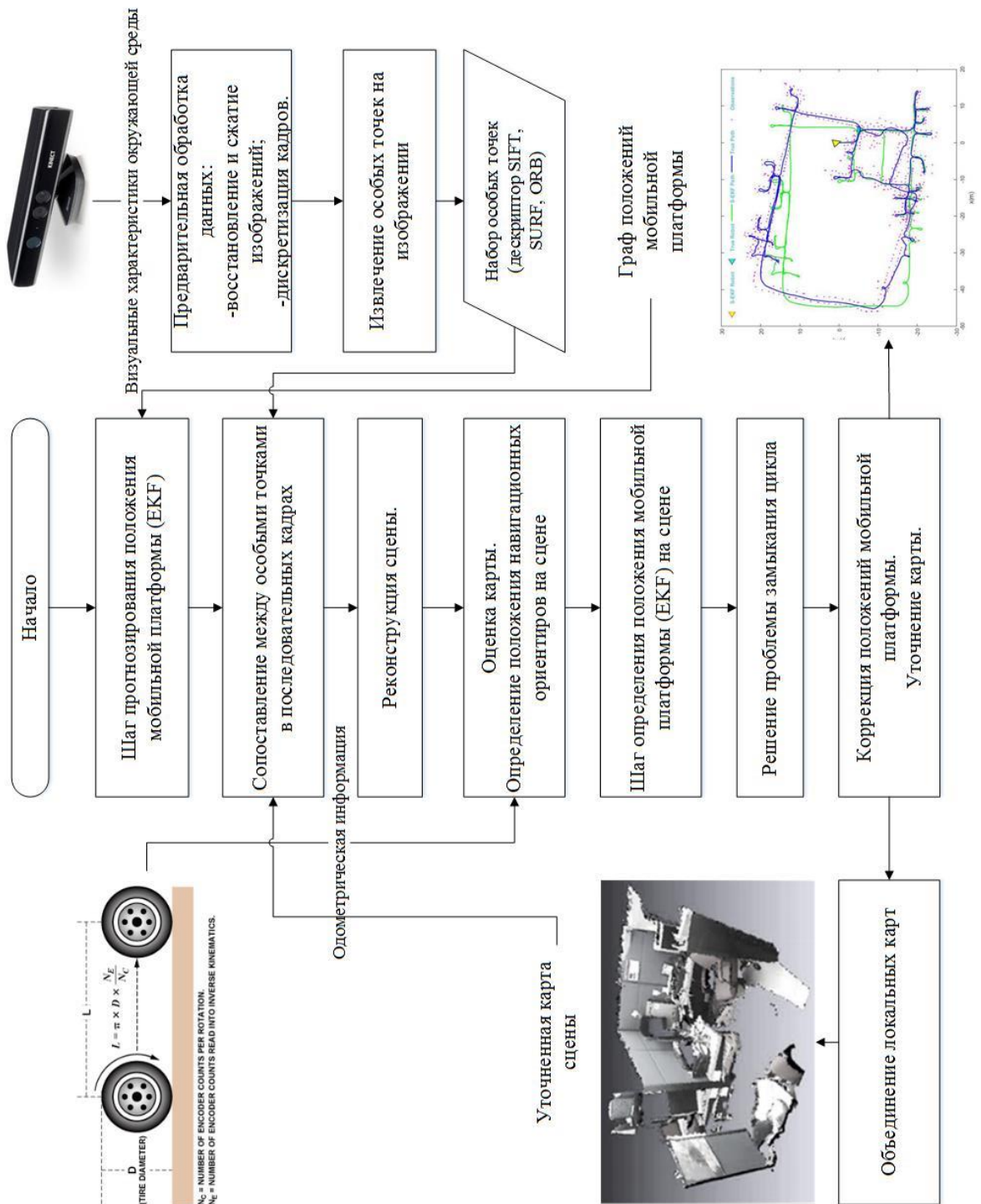


Рисунок 1.3 – Визуальная методология SLAM (Visual SLAM, RGB-D SLAM)

Шаг 8. Вычисление положения мобильной платформы на сцене с использованием ЕКФ: 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 9. Решение проблемы «замыкания цикла».

Шаг 10. Коррекция положения мобильной платформы на трехмерной сцене и уточнение карты.

Шаг 11. Объединение локальных карт. Изменение трехмерной карты окружающей среды, используя результаты, полученные на Шаге 9-10.

Важно отметить, что методы SLAM на основе визуальной методологии SLAM [54-56] могут сильно отличаться друг от друга, некоторые из них используют только визуальную информацию без одометрических данных, другие могут использовать при реконструкции сцены либо только данные о цвете (особые точки), либо только данные о глубине (трехмерные облака точек, карты глубины).

Методологиям Visual SLAM и EKF SLAM свойственны следующие недостатки:

1. 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.

2. Методологии имеют ограничения по применению в окружающей среде, искаженной различными естественными помехами, в том числе классическая методология Visual SLAM [54] показывает неудовлетворительные результаты для сцен с неравномерным освещением.

3. 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 (сцены), в следствие чего точность получаемых оценок положения мобильной платформы на карте относительно навигационных ориентиров не высока.

4. В методологии Visual SLAM информация о визуально связанных характеристиках, 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.

5. Визуальные методы для решения проблемы SLAM, которые основаны на итеративном алгоритме ICP для группы аффинных преобразований, не позволяют проводить точную регистрацию структурных элементов сцены.

6. 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. Такой подход при решении задачи регистрации данных приводит к снижению точности сопоставления трехмерных или двумерных данных, накладывает ограничения на сходимость (метод требует большего количества итераций, либо не сходиться совсем), приводит к снижению производительности всей итерационной процедуры.

7. 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.

Таким образом, для преодоления указанных проблем предлагается разработать новую методологию решения проблемы одновременной навигации и построения карты на основе комбинирования визуальных и семантических характеристик окружающей среды. Далее этой в главе рассмотрим более подробно некоторые аспекты решения задачи автоматической трехмерной реконструкции сцен и задачи планирования траектории применительно к решению поставленной в работе практической задачи спасения людей. Для этого будут рассмотрены некоторые аспекты решения задачи некооперативного распознавания личности по лицу на расстоянии. В последнем параграфе будут приведены результаты патентного поиска на основе базы данных ЦИТИС.

1.2 Автоматическая трехмерная реконструкция сцен

Автоматическая трехмерная реконструкция сцен окружающего пространства является одной из классических задач технического зрения современных роботизированных систем. Одновременная навигация и картографирование (SLAM) [2, 89] относятся к проблеме, в которой изображения (карты глубины) являются единственным внешним источником данных об окружающем пространстве и используются для решения двух взаимосвязанных задач: определение положения роботизированной системы в пространстве во время ее движения и создание трехмерной карты, которая соответствует зоне пространства, которая была исследована роботизированной системой.

Данная проблема имеет множество успешных решений для реконструкции небольших статических сцен с использованием датчиков дальности, таких как лазеры или гидролокаторы. Однако для динамичных, контекстуально сложных и крупномасштабных сцен построение точной трехмерной карты окружающего пространства является активной областью научных исследований. Для автоматической реконструкции качественных трехмерных моделей необходимо совместное использование данных, полученных с разных датчиков в каждый момент времени. Такой подход к решению задачи трехмерной реконструкции сцены обусловлен стремительным развитием различных типов датчиков, таких как камеры глубины, тепловизионные камеры, времяпролетные камеры и камеры видимого диапазона. На современном этапе при решении поставленной задачи большее распространение получили лазерные сканеры (времяпролетные камеры) и сенсоры глубины [90, 91]. Несмотря на то, что качество карты глубины в целом хорошее, существуют и проблемы с выходной информацией. Так в выходных данных образуются области неопределенности из-за того, что структурированный свет излучения после отражения не попадает на сенсор или быстрое движение камеры также приводит к потере данных. Современные подходы для реконструкции трехмерных моделей сцен с высокой точностью основаны на

совместном использовании данных, которые получают от различных типов мультисенсорных датчиков. В исследованиях часто используют комбинирование данных о характерных точках на изображениях, данные о положении датчика на трехмерной сцене и данные глубины. На сегодняшний день принята классификация методов трехмерной реконструкции сцены на два класса [92-94].

Методы первого класса используют для реконструкции трехмерной сцены различные двумерные проекции сцены из последовательности изображений (кадров). Методы данного класса при сопоставлении данных в двумерном пространстве используют информацию о характерных точках на изображении в виде дескрипторов (например, SIFT [95], SURF [96], ORB [97]). Каждый из перечисленных дескрипторов выполняет две процедуры: извлечение особых точек из изображения и создание описания информации об особых точках в определенном формате, с которым далее можно работать в других методах и алгоритмах. Во многих дескрипторах такое представление инвариантно к повороту и изменению масштаба.

Для сопоставления особых точек в двух последовательных изображениях необходимо вычислить расстояние между ними по информации в дескрипторе, при этом необходимо определить для всех точек в первом изображении ближайшие к ним соответствующие точки во втором изображении. В литературе по цифровой обработке информации представлено много способов определения ближайших точек [73], наиболее популярные методы из них: метод ближайших соседей в комбинации с алгоритмом k-d дерево или кэширование ближайших точек, поиск с использованием метода обратной калибровки, поиск ближайших соседей с ограничением в виде треугольника или эллипса (сферы). При выполнении процедуры сопоставления точек на изображениях для повышения точности метода целесообразно делать проверку обратного условия: для каждой точки второго изображения находить ближайшие к ней в первом изображении. Перед процессом реконструкции трехмерной сцены по ее проекциям необходимо удалить из множества соответствующих точек ошибочные соответствия (пары особых точек),

для это в области цифровой обработки информации применяют алгоритм устойчивой оценки параметров (RANdom SAmple Consensus, RANSAC) [98].

Методы первого класса характеризуют следующие главные недостатки: большая вычислительная сложность и требование наличия большого количества общих характерных точек (визуально связанных характеристик окружающей среды) в сопоставляемых кадрах для точной реконструкции сцены. Также следует отметить высокую чувствительность методов сопоставления особых точек к различного рода помехам на изображениях, которые могут быть вызваны наличием шумов (например, матричный шум приемников), искажением в виде размытия по причине движения датчика, неравномерным освещением. Такие помехи в комплексе с независимым определением особых точек в последовательности изображений приводят к снижению точности методов реконструкции трехмерной сцены по проекциям изображений.

Методы второго класса используются для анализа и обработки информации о глубине трехмерной сцены в виде трехмерного облака точек или карты глубины. Данные методы базируются на итеративном алгоритме ближайших точек (Iterative Closest Point, ICP) [71] или аналогах и используют для реконструкции трехмерной сцены инкрементальную вычислительную процедуру, связанную с решением задачи минимизации функции ошибки [75]. Таким образом, решение проблемы реконструкции трехмерной окружающей среды в методах данного класса основано на процессе сопоставления двух облаков точек с произвольным пространственным расположением. В алгоритме ICP на выходе алгоритма получается разреженное трехмерное облако точек, которое формируется на основе итеративного процесса вычисления геометрической структуры из перемещений датчика (мобильной платформы) по последовательности изображений (RGB-D кадров). При этом параметры датчика и координаты точек трехмерного облака уточняются на основе метод связок [75]. Таким образом, процесс сопоставления 3D облаков точек можно представить в виде двух шагов: первоначальное сопоставление данных и последующее уточнение процедуры регистрации. Главным недостатком методов

второго класса является необходимость выбора значений инициализации алгоритма пользователем при первом шаге работы алгоритма.

На ключевом шаге итеративного алгоритма ближайших точек решается задача нахождения такого геометрического преобразования между парой облаков трехмерных точек, которое позволяет совместить наилучшим образом эти облака точек на основе некоторой метрики близости. Рассматривают два типа геометрических преобразований: аффинное и ортогональное. Данный этап метода ICP часто называют вариационной подзадачей алгоритма ICP в варианте точка-точка, эквивалентной задачей для которой является задача абсолютной ориентации, известная проблема фотограмметрии [99, 100]. Точность реконструкции трехмерной сцены существенно зависит от способа решения вариационной задачи и выбора метрики для оценки соответствующего геометрического преобразования. Классический алгоритм ICP, предложенный Беслом П. [71], использует метрику точка-точка при минимизации функционала и может решаться с использованием метода оптимизации Левенберга-Марквардта, который обладает характерными для итеративных методов недостатками. Известно, что точность и сходимость итерационных методов при решении вариационной задачи, связанной с минимизацией функции ошибки, зависит от выбора начальных значений параметров метода ICP. Как известно, итеративный алгоритм ближайших точек инициализируется начальными значениями параметров геометрического преобразования, такими как поворот, параллельный перенос и изменение масштаба. Альтернативным способом решения вариационной задачи точка-точка являются точные аналитические решения, так называемые решения в замкнутой форме [99-101]. Исторически первыми были предложены решения вариационной задачи ICP в замкнутой форме для группы аффинных преобразований в работе Топмпсона Э. [99].

В работе Ду С. [102] исследована вариационная задача ICP относительно трех геометрических преобразований: параллельный перенос, поворот и изменение масштаба, при этом масштабирование может выполняться по трем осям. В

последнее время в работах Ду С. [103, 104] и Кобера В.И., Маковецкого А.Ю. [105] были предложены решения вариационной задачи ICP для группы произвольных аффинных преобразований. Кроме того, в работе [105] представлено решение вариационной задачи в замкнутой форме с метрикой точка-точка для ряда частных случаев, когда трехмерное облако точек на сцене является вырожденным. Данные результаты очень важны для проблемы одновременной навигации и картографирования, где при реконструкции окружающей среды приходится иметь дело с обработкой структурных элементов сцены. Хорн Б. предложил в своих работах решения вариационной задачи ICP в замкнутой форме для группы ортогональных преобразований с использованием кватернионов [100] и ортогональных матриц [101]. Кватернионы представляют из себя гиперкомплексные числа специального вида и являются удобным математическим аппаратом для обозначения операций положения и геометрического преобразования (поворот). На основе результатов, полученных в работах Хорна Б., был определен итеративный алгоритм ближайших точек, использующий метрику точка-точка [71].

Альтернативный вариант алгоритма ICP точка-точка сформулирован в работе Лоу К.-Л. [106]. Как было отмечено ранее, точность и сходимость метода ICP зависит от выбора метрики для решения вариационной задачи, наибольшее распространение при трехмерной реконструкции сцен получили метрика точка-точка (англ. наименование point-to-point) и метрика точка-плоскость (англ. наименование point-to-plane). Отличие методов ICP, использующих данные метрики для минимизации функционала, заключается в том, что в методе с метрикой точка-плоскость минимизируются расстояния от точек шаблона до касательных плоскостей к соответствующим точкам целевой поверхности, а в методе с метрикой точка-точка минимизируются расстояния между соответствующими точками.

В работе Русинкевича Ш. и Левоу М. [73] показано, что во многих приложениях, особенно при обработке больших объемов данных, применение метрики точка-плоскость при решении вариационной задачи ICP позволяет

добиться как лучшей сходимости, так и более высокой точности при сопоставлении облаков точек.

Аналогичным образом для группы ортогональных преобразований при решении вариационной задачи точка-плоскость при минимизации функционала используют метод оптимизации Левенберга-Марквардта или метод условной линеаризации углов [74]. В работе Ван дер Санде К. [107] было предложено точное решение вариационной задачи ICP для группы аффинных преобразований с использованием метрики точка-плоскость. Соответствие между точками двух облаков определяется на основе поиска ближайшего соседа. Этот шаг алгоритма является наиболее ресурсоемким. Для ускорения вычислений используется метод k-d дерева [73, 74]. Метод поиска ближайшего соседа при применении к реальным данным во многих случаях не дает удовлетворительного результата. С теоретической точки зрения метод также имеет недостатки, например, все точки одного облака могут быть сопоставлены одной точке другого облака. Наиболее эффективные алгоритмы восстановления трехмерной поверхности окружающей среды при произвольном изменении масштаба используют графическую карту занятости [108] и октодеревья [109]. Оценка положения камеры производится также с использованием алгоритма ICP [54]. И, наконец, устойчивость метода реконструкции трехмерной сцены достигается глобальной оптимизацией графа положений камеры и с помощью решения проблемы «замыкания цикла» [110, 111].

1.3 Некооперативное дистанционное распознавание личности по лицу

Предлагаемая комбинированная методология для решения проблемы одновременной навигации и построения карты в неизвестном пространстве может найти применение в решении различных научно-технических задач. Результаты диссертационного исследования были апробированы при создании автоматической системы управления мобильной платформой для решения задачи спасения людей. Для решения поставленной задачи необходимо использовать методы для некооперативного распознавания личности по лицу. В данном параграфе будет представлен обзор современного состояния исследований в области дистанционного распознавания личности по лицу.

На сегодняшний день в связи со стремительным ростом городской инфраструктуры резко увеличился поток людей на объектах транспорта, в торговых, офисных и развлекательных комплексах, во время проведения спортивных и культурных мероприятий. Обеспечение безопасности данных объектов и людей становится приоритетной задачей национальной безопасности, при этом важно обеспечить своевременную реакцию на внештатную или опасную ситуацию. Для решения данной задачи на объектах транспорта и в системах контроля доступа хорошие результаты показывают биометрические системы идентификации личности, которые позволяют автоматизировать процесс обнаружения потенциально опасных персон в потоке людей. В процессе идентификации личности системы используют физиологические (папиллярный узор пальца, геометрические и топологические особенности ладони, геометрия лица, структура радужной оболочки глаза, данные о распределении температуры кожи лица) и поведенческие (подчерк, походка, голос и речь человека, авторская принадлежность текста) биометрические признаки.

Для аутентификации человека по физиологическим или поведенческим критериям используются современные биометрические методы, которые можно разделить на кооперативные и некооперативные [112-114]. «Методы

распознавания по лицу [115, 116] используют различные методы получения информации: из одного изображения, из видео, из трехмерного изображения, из объектов, просвеченных инфракрасным светом [117]. После того, как лицо локализовано, применяются методы, основанные на внешнем виде лица или на геометрии лица [118]. Техники распознавания личности по лицу можно поделить на глобальные [119] и локальные [120], различают методы обработки, использующие двумерные и трехмерные модели лица [121, 122]. Существующие методы распознавания личности по лицу можно поделить на статистические методы и параметрические модели [123], метод главных компонент (Principal component analysis, PCA) [124], дискриминантный анализ, например, (Linear Discriminant Analysis, LDA), методы на основе эластичных графов, локальные бинарные шаблоны (Local Binary Patterns, LBP) [125], методы на основе 3D-дескрипторов поверхности лица [126, 127], которые основаны либо на определении средней и гауссовой кривизны, либо на определении геометрических отношений между особыми точками соответствующих поверхностей» [10].

Используются методы распознавания лица, инвариантные к мимике человека [128], также есть методы 3D обработки, позволяющие уменьшить влияние таких артефактов как различные прически, очки, усы. Кроме того, ведутся работы по применению динамической информации (изменение мимики лица [129], анализ походки и жестов [130, 131]) для решения задач распознавания и классификации.

Классический подход к решению задачи слежения за объектами на динамических сценах основан на использовании фильтрации Калмана [132] для отслеживания пользовательских поз при интерактивном компьютерном общении. Однако наилучшие результаты слежения с помощью фильтра Калмана достигаются при гауссовом распределении входных данных [133]. Идеи фильтрации Калмана, Марковские цепи и метод Монте-Карло использовались активно для создания новых алгоритмов слежения частиц [134-136], однако данные методы обладают высокой вычислительной сложностью и имеют проблемы со сходимостью алгоритмов [137]. Другие стратегии слежения используют: несколько гипотез [138, 139], ядро слежения [140], оптический поток [141, 142]. Перечисленные алгоритмы

слежения сталкиваются с серьезными проблемами, если объект слежения частично закрыт другими объектами, временно исчезает из кадра и потом появляется в кадре в другом месте, а также при наличии сильных фоновых помех. В последнее время были предложены методы, выделяющие различные характеристики из объекта и использующие их для поиска на динамических изображениях [143]. Был предложен быстрый адаптивный алгоритм, который учитывает ошибки слежения на предыдущих шагах и перестраивается таким образом, чтобы избежать ошибок слежения на последующих шагах [144]. Альтернативой традиционным методам слежения являются методы адаптивной линейной фильтрации [145-147] и система слежения за несколькими объектами на основе кинематической модели движения [148, 149].

Точность обнаружения человека в современных биометрических системах существенно зависит от условий проведения регистрации [150, 151]: освещенность, разрешение камер, угол обзора, структурированность потока и т.д. Кроме того, до сих пор не предложено методов, позволяющих контролировать значение ошибок распознавания при дистанционной идентификации личности. Таким образом, современные биометрические системы пассивной и кооперативной идентификации по лицу могут давать хорошие результаты только на близком расстоянии в контролируемых условиях, что является недостаточным для обеспечения безопасности объекта наблюдения в целом. Поэтому для решения поставленной задачи предлагаемые методы идентификации личности должны использовать при анализе множество биометрических характеристик и работать в условиях искажений окружающей среды и неструктурированного потока людей.

Среди основных научно-исследовательских работ в области идентификации личности по мультисенсорным биометрическим данным можно выделить следующие:

- Recognition of Noncooperative Individuals Using 3-D Face Biometric (University of Southern California) [154];
- Multimodal Biometric Fusion With Predictive Quality Metrics (General Electric Company) [155];

- Sensors, Surveillance and Biometrics Center of Excellence, Mantech Advanced Systems International, (Department of Defense Research and Development Laboratory, Massachusetts Institute of Technology (MIT) Lincoln Laboratory) [156];
- Image Stabilized Binoculars With Integrated 3-D Facial Recognition Imaging Capabilities (StereoVision Imaging, Inc) [157];
- Face Recognition From Video (University of Notre Dame) [158];
- Facial and Iris-Based Biometric System (New Jersey Institute of Technology) [159];
- Efficient, Field-Optimized, Multimodal Biometric System (International Biometric Group) [160];
- Facenorm: Normalization Plug-In for Improved Face Recognition of Noncooperative Individuals (Carnegie Mellon University) [161];
- Vocord.ru (Вокорд – системы биометрии и видеонаблюдения) [162].

1.4 Результаты патентного поиска

В работе был произведен патентный поиск по базе ЦИТИС <http://rosrid.ru/> по данной тематике:

1. АААА-А17-117030110075-0 Создание прототипа мобильной платформы с системами технического зрения, навигации и дополненной реальности.
2. АААА-А15-115122210112-5 Исследование научно-технических решений и разработка экспериментального образца многоцелевого двурукого робота-помощника на мобильной платформе, способного заменить человека при удаленном выполнении различных задач.
3. 01200853118 Разработка алгоритмов функционирования и конструкции мехатронного мини-модуля одноракурсного трехмерного технического зрения со структурированной подсветкой.
4. 116011510023 Развитие теории, методов и алгоритмов и создание технологии распознавания объектов, основанной на интеграции результатов обработки двумерных и трехмерных представлений сцены.
5. 01201159968 Разработка принципов построения, теоретических основ, программно-алгоритмических и аппаратных решений автоматизированного комплекса контроля кинематических параметров биомеханики позы и движений человека.
6. АААА-А17-117060110055-9 Исследование принципов построения систем технического зрения для мобильного наземного робототехнического комплекса.
7. 116012650002 Инструментальная среда для построения систем планирования движения в сложном трехмерном окружении.
8. АААА-А16-116060810022-5 Разработка новых представлений трехмерных неупорядоченных изменяющихся сцен в робототехнике,

поддерживающих эффективное обнаружение предметов, индексацию предметов, и пространственно-временной анализ на уровне предметов.

9. 01201151770 Разработка методов построения 3D-карты и системы управления движением мобильного робота на местности с произвольным рельефом.

10. 01201269303 Математические модели и морфологические алгоритмы для сравнения и классификации изображений в системах компьютерного зрения.

11. 114112570029 Проведение прикладных научных исследований для создания интеллектуальных технологий и программного обеспечения систем навигации и управления подвижными техническими средствами с применением методов машинного зрения и высокопроизводительных распределенных вычислений.

12. АААА-А17-117051210074-7 Исследование и разработка методов автономной калибровки и анализа положения конечностей антропоморфного робота на основе изображения, полученного с одной камеры.

13. АААА-А17-117101850050-5 Разработка и исследование алгоритмов технического зрения с использованием нейронных сетей для систем распознавания объектов на видеоизображениях.

14. АААА-А17-117101720041-3 Разработка программно-аппаратного комплекса обеспечения движения по гладкой и пересечённой местности антропоморфных робототехнических комплексов с приводами переменной жёсткости.

15. АААА-А17-117052310081-1 Исследование и разработка модели управления автономного роботизированного комплекса, предназначенного для эффективного и безопасного мониторинга техногенных образований.

Следует отметить наиболее близкие к предлагаемому в исследовании подходу отечественные разработки из патентной базы данных:

1. АААА-А17-117030110075-0 Создание прототипа мобильной платформы с системами технического зрения, навигации и дополненной реальности (Щеколдин А.И.).

2. АААА-А17-117060110055-9 Исследование принципов построения систем технического зрения для мобильного наземного робототехнического комплекса (Рудской А.И., Соснов Е.Н.).

3. 01201151770 Разработка методов построения 3D-карты и системы управления движением мобильного робота на местности с произвольным рельефом (Александров А.А., Зенкевич С.Л.).

4. АААА-А16-116060810022-5 Разработка новых представлений трехмерных неупорядоченных изменяющихся сцен в робототехнике, поддерживающих эффективное обнаружение предметов, индексацию предметов, и пространственно-временной анализ на уровне предметов (Лемпицкий В. С.).

Целью первого проекта из вышеприведенного списка является обучение и вовлечение студентов в разработку новых робототехнических систем, привитие им навыков работы в областях систем автоматического регулирования, навигации и систем дополненной реальности. Целью второго проекта является разработка концепции системы машинного зрения робототехнической системы, которая позволяет вычислять информацию о положении объектов на трехмерной сцене в реальном масштабе времени. Как отмечают авторы проекта, основная проблема современных систем очувствления мобильной робототехники связана с большим объемом вычислений, необходимых для восстановления трехмерного образа окружающей обстановки. Поэтому в проекте не производится точная трехмерная реконструкция объектов и окружающей среды, что серьезно ограничивает функциональность данного робототехнического комплекса. Целью третьего проекта является разработка алгоритмов и программных средств, обеспечивающих движение мобильного робота по местности с произвольным рельефом и недетерминированными препятствиями в режиме супервизорного управления или в автономном режиме. Информация о методах и алгоритмах трехмерной реконструкции объектов и окружающей среды отсутствует. Более того, так как проект был выполнен в 2011 году, новые эффективные алгоритмы, такие как не итеративные алгоритмы ICP для сопоставления облаков точек точка-плоскость, не могли быть использованы при выполнении проекта, так как они появились в печати

2016-2017 годах. Целью четвертого проекта является построение методов трехмерного представлений сцен, которые поддерживали бы нахождение (локализацию) предметов, заданных с помощью двумерных фотографий. Таким образом, основная идея – упростить трехмерную модель окружающей среды (использовать двумерные проекции) для уменьшения вычислительных затрат, упрощения принятия решений робототехническим комплексом. Принципиальное отличие предлагаемого решения от существующих заключается в использовании коллективного технического зрения группы мобильных платформ для построения точной трехмерной карты окружающей среды, а также согласованного использования данной информации для эффективного решения различных прикладных задач. Ожидается, что новые методы и алгоритмы позволят создать прототип системы на основе группы мобильных платформ, решающий широкий спектр роботизированных задач, с лучшим качеством, чем известные существующие системы, работающие в аналогичных условиях. Нас сегодняшний день известен 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 рамках доклада «Фонд перспективных исследований в системе оборонных инноваций» общественного совета председателя военно-промышленной комиссии при правительстве РФ (протокол №1/2013) были определены приоритетные направления исследований и разработок для Вооруженных сил РФ, среди которых следует выделить приоритетное направление исследований и разработок: военная робототехника. Следует отметить следующие научно-исследовательские проекты в области SLAM:

– Stanley (DARPA, Stanford Artificial Intelligence Lab, the Stanford Racing Team, Volkswagen Electronics Research Laboratory) [163];

- Robots Podcast: 3D SLAM (Department of Defense Research and Development Laboratory, Massachusetts Institute of Technology (MIT) Lincoln Laboratory) [164];
- Autonomous and Perceptive Systems research (University of Groningen) [165];
- Active 3-D Visual SLAM (General Electric Company) [166];
- SLAM using vision (Imperial College London) [167];
- WPI Robotics Engineering C Squad (Worcester Polytechnic Institute) [168];
- JPL / UCSB / Caltech (Jet Propulsion Laboratory, Pasadena) [169];
- TRAC Labs (TRAC Labs Inc., Webster) [170];
- Volkswagen of America (GE, Electronics Research Laboratory) [171];
- Team IHMC (Institute for Human and Machine Cognition) [172];
- TORC (TU Darmstadt, Virginia Tech) [173].

Известные открытые системы SLAM HOGMan, ORB-SLAM, TOROSLAM, RGBDSLam, EKFMonoSLAM используются в перечисленных коммерческих системах: Stanley, 3D SLAM, Autonomous and Perceptive Systems research и Active 3-D Visual SLAM. 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. Можно отметить следующие решения: YUMI и Roberta (ABB), APAS (Bosh), CR-35iA (Fanuc), PRob 1R (F&P Personal robotics), NEXTAGE, IIWA (KUKA), Speedy-10 и Max-150 (Mabi). Коллоборативный робот будущего – это полностью автономный антропоморфный робот с собственной системой навигации в пространстве, системами трехмерного зрения, понимания и генерации речи, системой поддержки принятия решений и прикладным функционалом, связанным с его сферой деятельности.

1.5 Выводы по главе

1. Методологиям Visual SLAM и EKF SLAM свойственны следующие недостатки:

- 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;

- 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;

- 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 (сцены), в следствие чего точность получаемых оценок положения мобильной платформы на карте относительно навигационных ориентиров не высока;

- в методологии Visual SLAM информация о визуально связанных характеристиках, 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. Такой

подход при решении задачи регистрации данных приводит к снижению точности сопоставления трехмерных или двумерных данных, накладывает ограничения на сходимость (метод требует большего количества итераций, либо не сходится совсем), приводит к снижению производительности всей итерационной процедуры;

– 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.

2. Для решения задачи автоматической реконструкции трехмерной сцены в работе будет предложен новый метод на основе итеративного алгоритма ближайших точек (Iterative Closest Point, ICP).

3. В работе проведен патентный поиск по базе данных ЦИТИС, представлены известные решения в области SLAM. Принципиальное отличие предлагаемой методологии от существующих подходов заключается в создании совместного комбинированного решения проблемы одновременной навигации и картографирования для группы мобильных платформ с целью построения точной трехмерной карты окружающей среды, а также согласованного использования данной информации для эффективного решения прикладных задач. Ожидается, что новые методы и алгоритмы позволят создать прототип системы на основе группы мобильных платформ, решающий широкий спектр роботизированных задач, с лучшим качеством, чем известные существующие системы, работающие в аналогичных условиях.

4. Основные результаты по данной главе опубликованы в статьях автора данного исследования в разделах введение в публикации [9-38].

ГЛАВА 2 КОМБИНИРОВАННАЯ МЕТОДОЛОГИЯ ДЛЯ РЕШЕНИЯ ПРОБЛЕМЫ ОДНОВРЕМЕННОЙ НАВИГАЦИИ И ПОСТРОЕНИЯ КАРТЫ

В последнее время было разработано множество методов для решения задачи локализации и навигации мобильных роботов в помещении на основе визуальной методологии, например, RGBD-Slam, ORB-SLAM, TreeMap, Sparse Least Squares on Manifolds, RobotVision, OpenSeqSLAM, FALKOLib. В данной главе предложена комбинированная методология для решения проблемы одновременной навигации и построения карты, использующая визуально связанные характеристики окружающей среды и информацию о глубине для сопоставления изображений, регистрации 3D облаков точек, а также построения глобальных последовательных 3D карт окружающего пространства.

Производительность и вычислительная сложность предлагаемой комбинированной методологии решения проблемы SLAM представлены и обсуждены на примере эталонных баз данных. Полученные научные и практические результаты исследования могут быть применены при решении различных практических задач:

1. Навигация мобильной платформы на сцене в реальном масштабе времени.
2. Планирование траектории движения мобильной платформы на сцене со статическими и динамическими препятствиями.
3. Слежение за динамическими объектами на сцене.
4. Реконструкция трехмерных сцен с семантической разметкой объектов на анализируемой сцене.

2.1 Постановка задачи. Современное состояние проблемы SLAM

Проблема одновременной навигации и построения карты (Simultaneous localization and mapping, SLAM) [1-8] заключается в реконструкции двухмерной или трехмерной карты неизвестного окружающего пространства вокруг автономного интеллектуального объекта управления, например, мобильной платформы и определении своего положения на этой карте в каждый момент во время движения. На сегодняшний день разработано множество методологий решения проблемы одновременной навигации и построения карты, самые известные из них:

1. Методология быстрого SLAM на основе частичного фильтра (Particle Filter SLAM). Также в литературе используется альтернативное название FAST SLAM [52].
2. Методология SLAM на основе расширенного фильтра Калмана (Extended Kalman Filter, EKF) [51, 53].
3. Визуальная методология SLAM (Visual SLAM) [54-56].
4. Методология SLAM на основе графов (Graph-Based SLAM) [57].

Методология на основе EKF SLAM аналогична известному алгоритму на основе расширенного фильтра Калмана [53], который используется в различных приложениях для локализации объекта управления в пространстве или в прогнозировании. Основное различие связано с размерностью вектора состояния, т.к. решение проблемы на основе EKF SLAM требует анализа больших объемов данных, в связи с чем вектор состояния имеет существенно большую размерность. В связи с этим решение поставленной проблемы на основе EKF SLAM имеет высокую алгоритмическую и вычислительную сложность.

Пусть t – шаг алгоритма, управляющие сигналы $u_t = \{u_1, u_2, \dots, u_t\}$, наблюдения $z_t = \{z_1, z_2, \dots, z_t\}$, карта сцены m , путь мобильной платформы $x_t = \{x_0, x_1, \dots, x_t\}$. Тогда представим графическую интерпретацию методологии на

основе EKF SLAM в виде следующей схемы (Рис. 2.1), которая представлена в работе Трана С. [53].

Методология FAST SLAM [51] имеет ограниченный набор приложений, в которых анализируются и обрабатываются сцены ограниченной размерности, для реконструкции крупномасштабных сцен и навигации на них данная методология не может быть применена. Основные недостатки методологии быстрого SLAM связаны с плохой масштабируемостью фильтра частиц и невозможностью выполнять точное сопоставление частиц в условиях неопределенности.

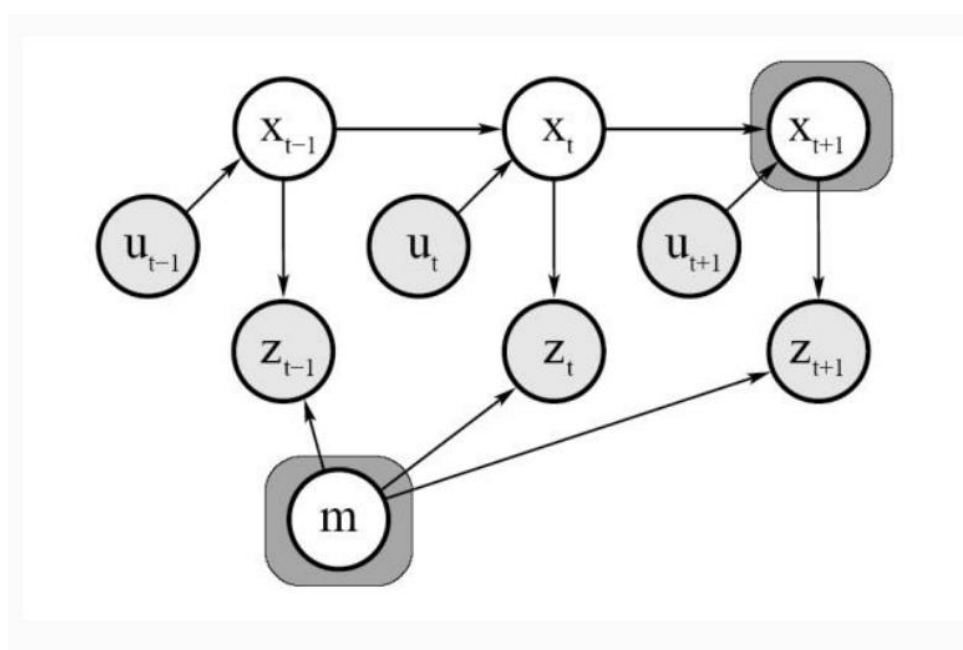


Рисунок 2.1 – Графическая интерпретация методологии SLAM на основе EKF

В Graph-Based SLAM [57] узел разреженного графа соответствует положению мобильной платформы, а связи в графе соответствуют относительным позициям внутри вектора, задающего положение платформы в пространстве [59, 60]. Современные методологии SLAM для крупномасштабных сцен основаны на визуальной методологии SLAM или методологии на основе EKF SLAM.

Как было сказано, проблема SLAM состоит из двух задач, которые решаются одновременно, одна из задач связана с построением карты окружающего

пространства. Современные методы для реконструкции трехмерных сцен представляют следующую последовательность шагов:

- извлечение информации о визуальнo связанных характеристиках окружающей среды (особых точках) из изображений;
- решение задачи регистрации для данных о цвете и для данных о глубине;
- решение задачи выравнивания в трехмерном пространстве последовательности RGB-кадров;
- глобальная оптимизация.

Во многих задачах робототехники при использовании датчиков глубины часто используется итеративный алгоритм ближайших точек (Iterative Closest Point, ICP), либо его многочисленные варианты [71-73]. Данный алгоритм позволяет минимизировать расстояние между двумя трехмерными геометрическим моделями, полученными с разных точек обзора трехмерной сцены, на основе некоторой метрики. Итеративный алгоритм ближайших точек может завершиться либо по достижении определенного порога точности, задаваемого пользователем, либо по одному или нескольким критериям остановки работы алгоритма, например, максимально допустимое число итераций. Результатом работы алгоритма ICP является геометрическое преобразование, которое в общем случае содержит матрицу поворота, матрицу масштаба и вектор параллельного переноса.

Для решения проблемы обнаружения цикла (петли) в графе современные методологии часто применяют подход на основе Bag of Words. После обнаружения петли в графе положений мобильной платформы полученное соответствие между несколькими последовательными RGB-D кадрами помечается в виде ограничения в графе положений. Рассмотренные выше способы применения визуальной методологии SLAM показывают хорошие результаты с точки зрения точности и сходимости в случае получения исходных данных для эксперимента от датчиков глубины. Ранее в параграфе 1.1 были рассмотрены недостатки известных методов регистрации данных на основе итеративного алгоритма ближайших точек. С целью преодоления названных недостатков методов ICP в данном исследовании будет

предложена новая комбинированная методология решения проблемы одновременной навигации и картографирования.

Для того чтобы автономные интеллектуальные системы могли взаимодействовать со своим окружением, они должны использовать при обработке данных как геометрические, так и семантические свойства окружающей их сцены. Большинство известных исследований до настоящего времени рассматривали эти проблемы одновременной навигации и картографирования отдельно, предлагая решение проблемы SLAM либо на основе геометрического, либо на основе семантического преобразования данных.

В предложенной методологии проблема SLAM включает как семантически значимые компоненты трехмерной сцены, так и их геометрические представления на основе особых точек (визуально связанные характеристики окружающей среды) и трехмерных облаков точек и/или карт глубины. Схема предлагаемой методологии представлена на рисунке 2.2.

При решении проблемы локализации в работе одновременно создаются геометрические модели облака точек для трехмерных объектов, соответствующим некоторым семантическим классам предметной области, и трехмерная карта сцены, которая содержит эти модели объектов в качестве объектов сцены: препятствий или семантических ориентиров.

Заметим, что автономные интеллектуальные системы (мобильные платформы) перед исследованием окружающего пространства не имеют информации о семантических свойствах анализируемой сцены и принадлежности объектов к конкретным классам предметной области. Предварительное знание о предметной области включается в методологию решения проблемы SLAM путем встраивания системы классификации объектов на основе известного детектора объектов YOLO (You Only Look Once).

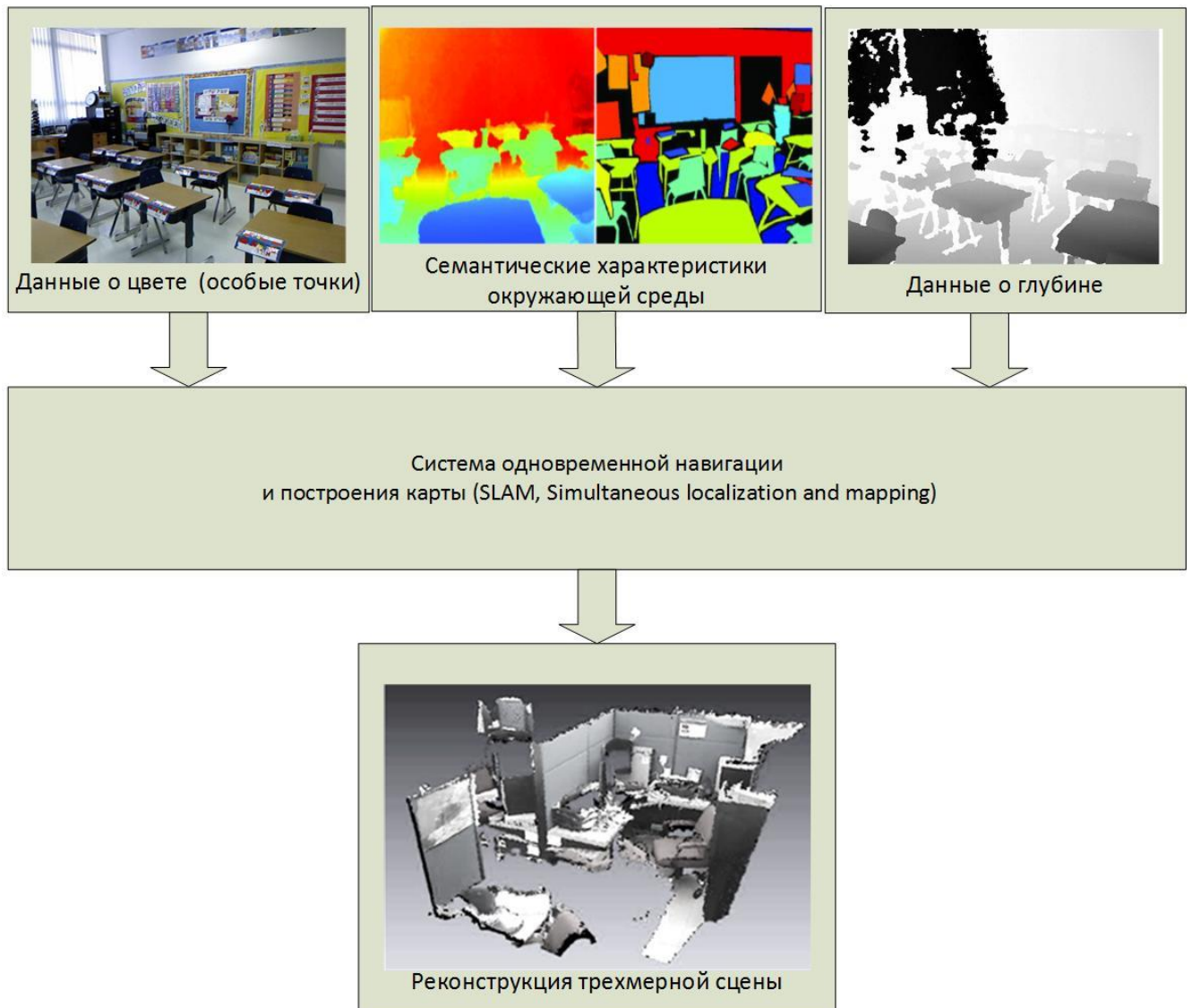


Рисунок 2.2 – Комбинированная методология решения проблемы SLAM: входные и выходные данные

2.2 Комбинированная методология при решении задачи регистрации данных

В комбинированной методологии входной информацией являются последовательность RGB-D кадров и критерии останова для алгоритмов регистрации и сопоставления данных. Входные данные для компьютерного моделирования и вычислительных экспериментов с реальными данными могут быть получены как с одной, так и с множества камер, расположенных на мобильной платформе. В качестве мультисенсорных датчиков могут использоваться камеры видимого диапазона, камеры глубины, RGB-D камеры. Каждый RGB-D кадр содержит данные двух типов: изображения (данные о цвете) и карта глубины (данные о глубине).

Выходной информацией являются реконструкция трехмерной карты сцены и граф положений мобильной платформы, связанный с координатами трехмерной карты в различные моменты времени движения мобильной платформы. В комбинированной методологии проблема SLAM может быть представлена в виде итерационной процедуры (Рис. 2.3), состоящей из следующей 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, вычисляемых в круглом скользящем окне (ДЛ ГНГ).

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

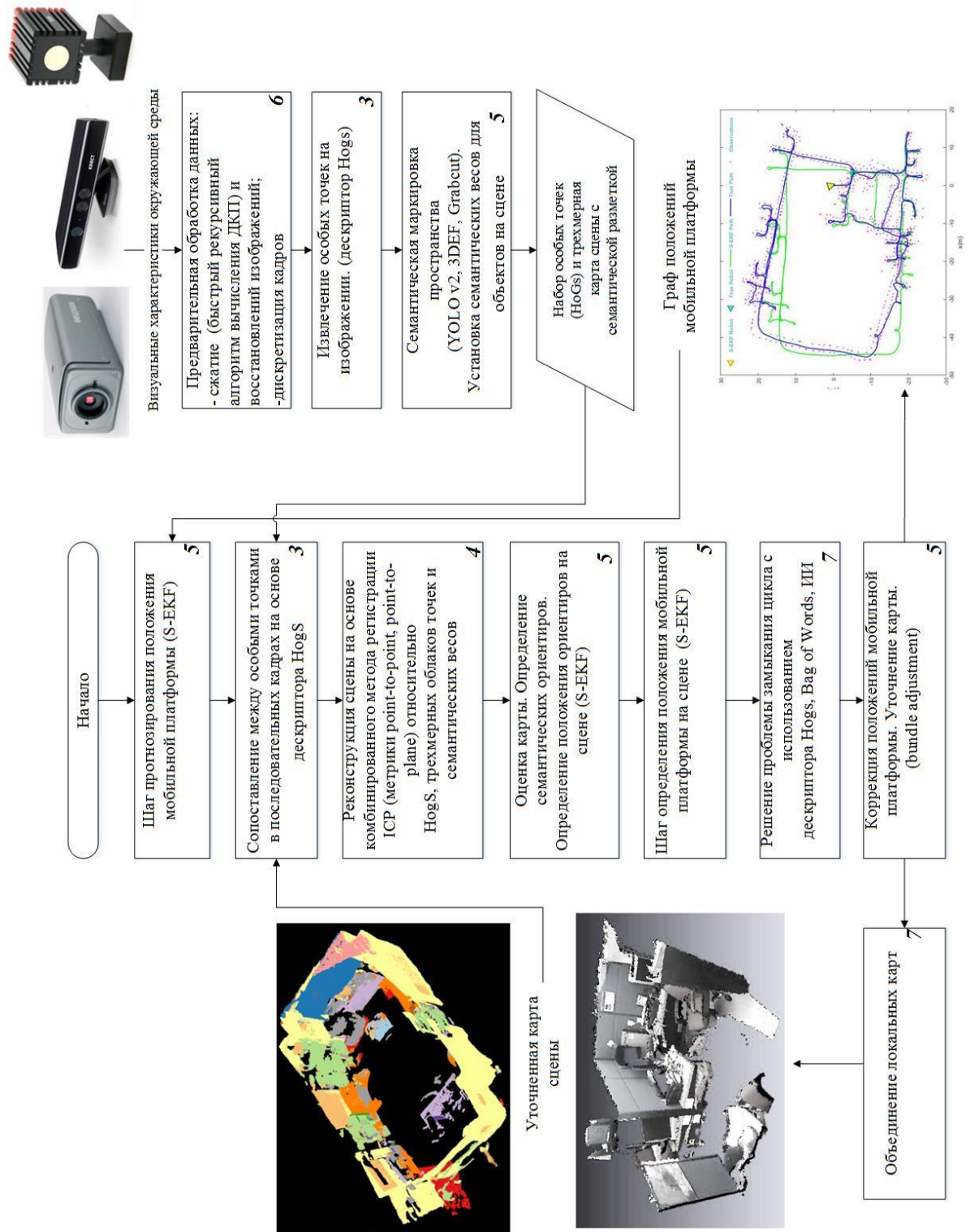


Рисунок 2.3 – Комбинированная методология SLAM

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. Двумерные данные (изображения) сопоставляются с использованием предложенного в работе алгоритма сопоставления данных на основе дескриптора ДЛ ГНГ. Трехмерные данные (облака точек) сопоставляются с использованием метода ближайших соседей.

Шаг 8. Реконструкция сцены на основе комбинированного метода ICP с метрикой точка-точка и метрикой точка-плоскость для группы аффинных и 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 ДЛ ГНГ, 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 мобильной платформы [98].

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

На рисунке 2.3 цифра в правом нижнем углу блока соответствует номеру главы в диссертационной работе и номеру задачи в диссертационной работе, в которой представлен соответствующий метод, алгоритм или подход.

Для цифровой обработки изображений используются дескрипторные методы, извлекающие из изображений особые точки (визуально связанные характеристики окружающей среды), которые впоследствии используются для сопоставления в последовательности RGB-D кадров (Рис. 2.4).

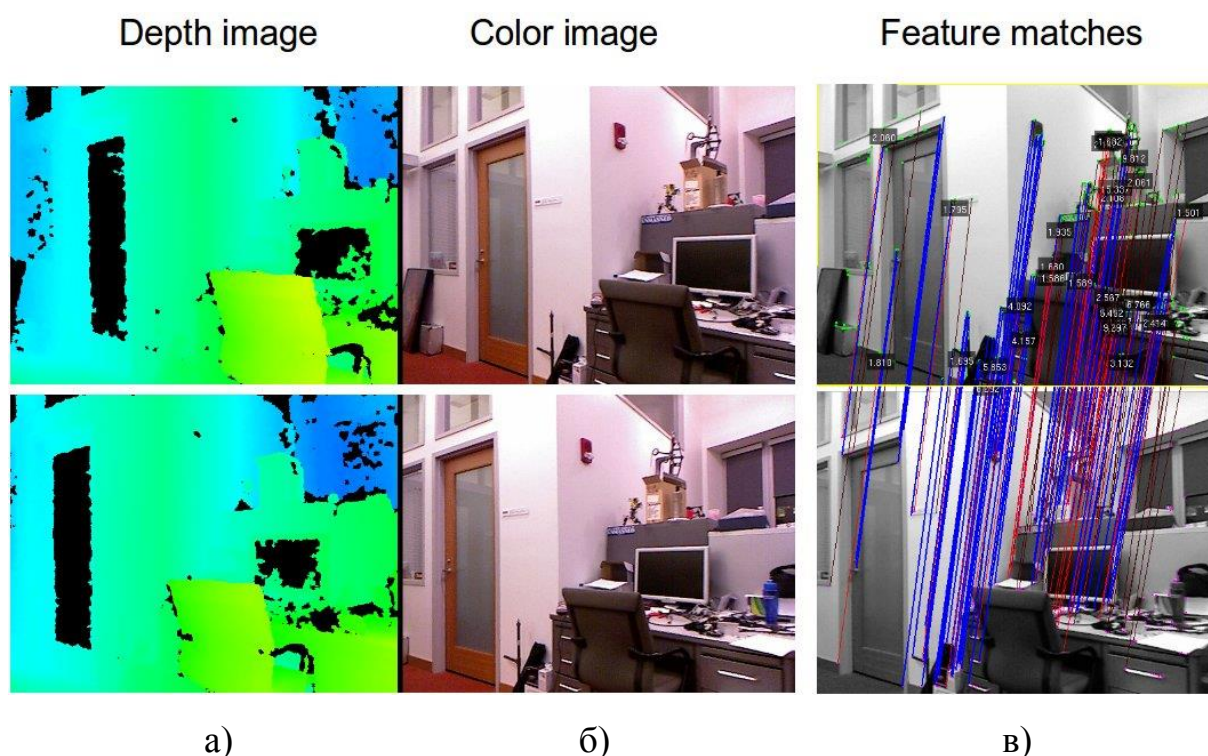


Рисунок 2.4 – Входные данные для компьютерного моделирования на примере эталонной базы данных: а) карта глубины, б) изображение, в) соответствие между особыми точками в двух последовательных RGB-D кадрах

Разработанная методология основана на использовании трех типов информации для решения проблемы SLAM: визуально связанные характеристики окружающей среды (изображения, данные о цвете), трехмерные облака точек (данные о глубине сцены), данные о семантических свойствах окружающей среды. Принципиально важно, что перечисленные источники информации используются на всех ключевых этапах решения проблемы SLAM (шаги 6, 8, 10, 11, 13), а не при решении отдельных задач.

Информация о семантических характеристиках окружающей среды в предлагаемой методологии получена в результате использования известных специализированных методов и алгоритмов YOLO и 3DEF (3D Entangled Forest), которые позволяют проводить семантическую маркировку пространства (Рис. 2.5) на основе данных, получаемых с различных мультисенсорных датчиков (RGB-камеры, камеры видимого диапазона). Данные методы позволяют решать

известную в робототехнике проблему привязки символов (семантических маркеров), которая заключается в том, что текущее состояние системы (мобильной платформы) в каждый момент времени задается не только сенсорными данными, но и данными о предметной области на сцене.

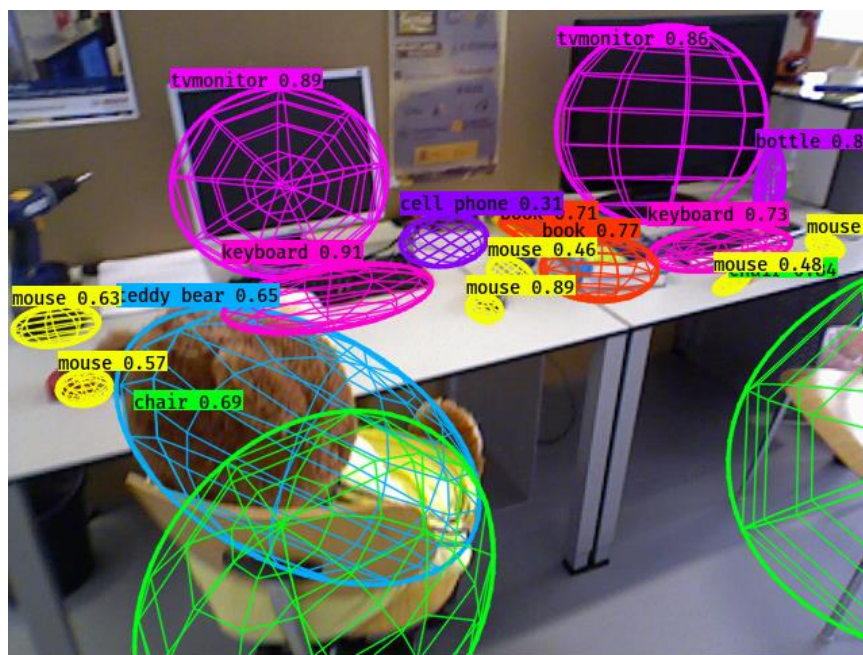


Рисунок 2.5 – Пример семантической маркировки окружающего пространства на основе изображений в RGB-D кадрах и дескриптора для классификации объектов (проект Quadric SLAM)

При семантической маркировке пространства объекты разных классов на анализируемом изображении выделяются разным цветом и сопровождаются атрибутом – вероятностью отнесения объекта к соответствующему классу предметной области. В предлагаемой методологии семантика концептуальных знаний мобильной платформы полностью основана на информации, которая получается с множества мультисенсорных датчиков. Таким образом, в качестве данных об окружающем пространстве используются данные мультисенсорных датчиков, которые осуществляют мониторинг окружающей среды и параметров мобильной платформы.

В результате постоянно повторяющейся последовательности шагов при комбинированном решении проблемы SLAM формируется множество абстрактных состояний и свойств окружающей среды, которые постоянно пополняют концептуальные знания мобильной платформы об окружающем мире. В работе для представления семантики используется символьное представление для множества понятий, которые были получены в результате процесса семантической маркировки пространства. Каждому объекту, который был идентифицирован на трехмерной сцене ставится в соответствие семантический маркер (ориентир) и его вес. Значения весов для классов предметной области определяются экспертами. Над понятиями, в более узком смысле объектами, можно совершать обычные операции вывода или использовать их в классических методах искусственного интеллекта как концепты (навигация, трехмерная реконструкция, планирование траектории). Для представления знаний в работе использована модель онтологий, разработанная автором [29, 35].

Комбинированная методология решения проблемы SLAM предполагает, что визуальная и семантическая информация о трехмерной сцене будет использоваться на всех ключевых шагах решения проблемы SLAM:

Шаг 1. Реконструкция трехмерной карты окружающей среды.

Шаг 2. Навигация и определение положения мобильной платформы на карте.

Шаг 3. Решение проблемы «замыкания цикла».

Шаг 4. Глобальная оптимизация, выравнивание последовательности RGB-D кадров.

Рассмотрим один из ключевых шагов данной методологии – Шаг 1, который связан с решением задачи регистрации данных. Данный шаг связан с комбинированным решением вариационной задачи алгоритма ICP на основе данных о визуальных и семантических характеристиках окружающей среды. Задача регистрации данных заключается в построении модели объекта по множеству его проекций в окружающем пространстве, которые были получены либо из разных положений датчиков вокруг камеры, либо за счет вращения объекта вокруг датчика. В данной работе используется первый подход (Рис. 2.6).

Рассмотрим частный случай решения вариационной задачи ICP и на основе анализа полученных данных о результатах компьютерного моделирования и вычислительного эксперимента на реальных данных сделаем выводы об достоинствах и недостатках предлагаемой комбинированной методологии SLAM.

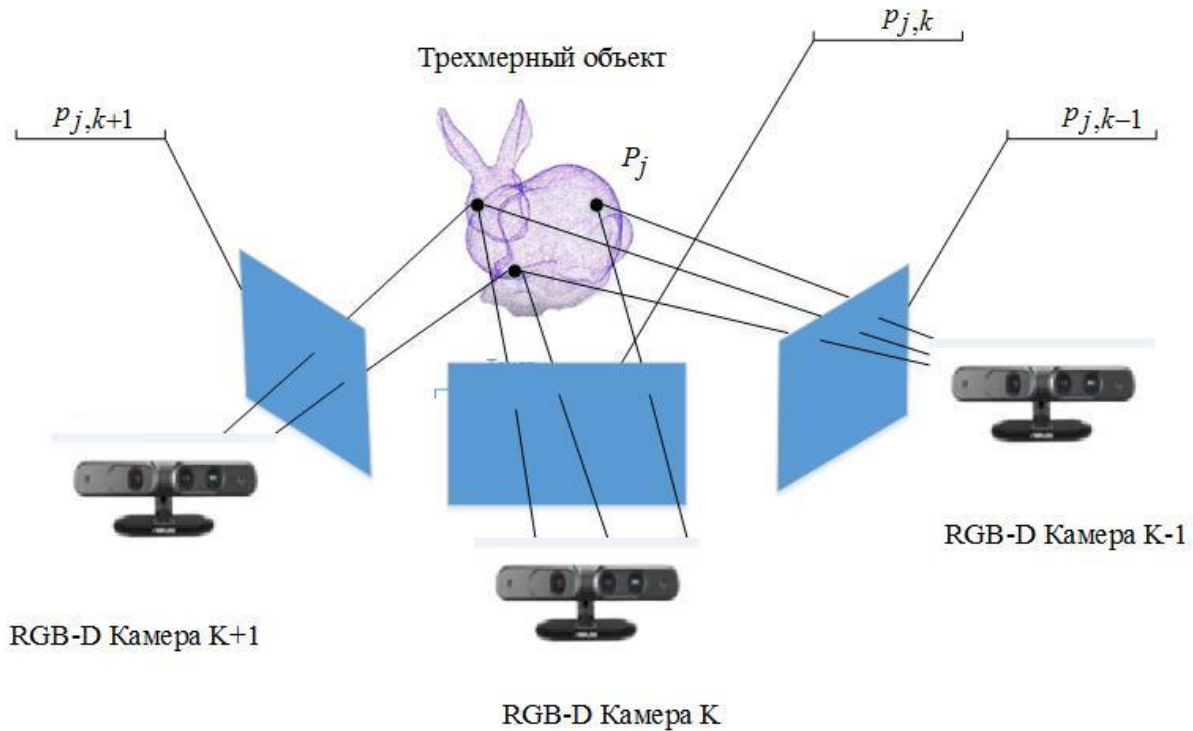


Рисунок 2.6 – Графическая интерпретация условий эксперимента метода регистрации данных для сцены со множеством RGB-D камер, k – номер камеры на сцене, P_j – точка в трехмерном облаке точек X (Y)

Введем следующие определения: $X = \{x_1, \dots, x_n\}$ – исходное трехмерное облако точек и $Y = \{y, \dots, y_m\}$ – целевое трехмерное облако точек в $\mathbb{R}^3$. В параграфе 2.2 было показано, что итеративный алгоритм ближайших точек (ICP) [71-73] может быть представлен как геометрическое преобразование между исходным облаком точек X в целевое облако точек Y следующего вида

$$Rx_i + T, \quad (2.1)$$

где R – матрица, содержащая компоненты преобразования поворот, T – вектор, содержащий компоненты преобразования перенос, $i = 1, \dots, n$.

Известны варианты алгоритма ICP, которые кроме преобразования поворота и переноса используют преобразование масштаба (S-ICP)

$$RSx_i + T, \quad (2.2)$$

где S – матрица, содержащая компоненты преобразования масштаба.

Входной информацией для алгоритма ICP являются: облака точек X и Y ; значение матрицы компонентов преобразования поворот R и значение вектора компонентов преобразования перенос T при инициализация алгоритма (первый шаг алгоритма ICP); набор значений параметров для остановки работы алгоритма (критерии останова). Выходной информацией для алгоритма ICP является такое геометрическое преобразование, которое минимизирует расстояние между облаками точек X и Y .

Представим классический метод ICP в виде следующей итерационной процедуры:

1. Предварительная обработка данных в одном или обоих плотных облаках точек X и Y с целью уменьшения количества точек в анализируемых данных.

2. Установление соответствия между точками в X и Y : определение соответствующих точек, таких что $\forall x_k \in X, y_k = \{y \in Y | y(x_k, y) = \min\}$. Существуют несколько способов определения соответствия точек: поиск ближайшей точки в другом облаке точек (вычисление может быть ускорено с помощью k-d дерева или кэширования ближайших точек); поиск пересечения лучей, исходящих из точки исходного облака точек в направлении нормали исходной точки к поверхности целевого облака точек; поиск с использованием метода обратной калибровки (проекция точки из исходного облака точек на целевое облако точек с выполнением поиска в изображении диапазона назначения).

3. Процедура взвешивания определенных пар соответствующих точек с использованием весовых коэффициентов. Существуют разные стратегии назначения весов: назначение постоянных весов; присвоение меньших весов парам

с большим межточечным отношением; взвешивание на основе совместимости нормалей.

4. Процедура отклонения определенных пар соответствующих точек для исключения выбросов в данных, при этом в процессе отбора пар может рассматриваться каждая пара точек в отдельности или анализироваться полный набор пар в двух облаках точек X и Y .

5. Назначение метрики качества классификации для соответствующих пар точек. Существуют разные типы метрик, наиболее часто при решении вариационной задачи ИСР используются метрика точка-точка и метрика точка-плоскость.

6. Ключевой этап алгоритма ИСР: нахождение решения вариационной задачи с использованием выбранной на шаге 5 метрики. Различают несколько вариантов минимизации функции ошибки: многократное создание набора соответствующих точек с использованием текущего преобразования и поиск нового преобразования; итеративная минимизация в сочетании с экстраполяцией в пространстве преобразования; итеративная минимизация с использованием различных произвольно выбранных подмножеств точек и выбор оптимального преобразования с использованием метрики наименьшей медианы квадратов; стохастический поиск наилучшего преобразования с использованием метода имитации отжига.

Можно представить функцию минимизации ошибки для определения пар совпадающих точек в следующем виде

$$E(R, S, T) = \frac{1}{N} \sum_{i=1}^N \|(RSx_i + T) - y_i\|^2. \quad (2.3)$$

Пусть lm – количество соответствующих пар точек. Давайте найдем минимум новой функции сопоставления кадров. Если выражение (2.3) имеет минимум, то оно удовлетворяет следующему уравнению

$$\frac{dE(T)}{dT} = \frac{2S}{N} \sum_{i=1}^N ((RSx_i + T) - y_i) = 0. \quad (2.4)$$

Из (2.4) следует, что $T = \frac{1}{N} \sum_{i=1}^N y_i - \frac{1}{N} \sum_{i=1}^N RSx_i$. Таким образом, для того чтобы достигнуть минимума, функция сопоставления должна быть представлена как

$$E(R, S) = \frac{1}{N} \sum_{j=1}^N \left\| RS \left(x_i - \frac{1}{N} \sum_{i=1}^N x_i \right) - \left(y_i - \frac{1}{N} \sum_{i=1}^N y_i \right) \right\|^2. \quad (2.5)$$

Пусть $a_i = \left(x_i - \frac{1}{N} \sum_{i=1}^N x_i \right)$, $b_i = \left(y_i - \frac{1}{N} \sum_{i=1}^N y_i \right)$, тогда функция минимизации ошибки (функция сопоставления) может быть представлена как

$$E(R, S) = \frac{1}{N} \sum_{i=1}^N \|RSa_i - b_i\|^2. \quad (2.6)$$

Затем

$$E(R, S) = \frac{1}{l_m} \left(\sum_{i=1}^{l_m} a_i^T S^2 a_i - 2 \sum_{i=1}^{l_m} b_i^T R S a_i + \sum_{i=1}^{l_m} b_i^T b_i \right). \quad (2.7)$$

Для того чтобы найти минимум данной функции сопоставления, необходимо найти решение следующего дифференциального уравнения:

$$\frac{dE(R, S)}{dR} = 0. \quad (2.8)$$

Чтобы получить матрицу масштаба из уравнения (2.7), сделаем следующие преобразования

$$\frac{dE(R, S)}{dS} = 2 \sum_{i=1}^{l_m} a_i^T S E_j a_i - 2 \sum_{i=1}^{l_m} b_i^T R S E_j a_i = 0, \quad (2.9)$$

где $E_j = (0, \dots, 0, 1, 0, \dots, 0)$, $j = 1, 2, 3$ есть диагональная матрица, j есть скаляр, который определяет размерность матрицы. Тогда определим значение элемента матрицы масштаба следующим образом

$$S_j = \frac{\sum_{i=1}^{l_m} b_i^T R E_j a_i}{\sum_{i=1}^{l_m} a_i^T R E_j a_i}. \quad (2.10)$$

Рассмотрим решение вариационной задачи итеративного алгоритма ближайших точек [20, 32]

$$J(A, T) = \sum_{i=1}^n \| A x_i + T - y_i \|^2. \quad (2.11)$$

Тогда вариационная задача ICP примет вид

$$\underset{A, T}{\operatorname{argmin}} J(A, T), \quad (2.12)$$

где

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, T = \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix}, x_i = \begin{pmatrix} x_{1i} \\ x_{2i} \\ x_{3i} \end{pmatrix}, y_i = \begin{pmatrix} y_{1i} \\ y_{2i} \\ y_{3i} \end{pmatrix}. \quad (2.13)$$

Можно заметить, что

$$\begin{aligned} A = J(A, T) = \sum_{i=1}^n & (a_{11}x_{1i} + a_{12}x_{2i} + a_{13}x_{3i} + t_1 - y_{1i})^2 + \\ & + (a_{21}x_{1i} + a_{22}x_{2i} + a_{23}x_{3i} + t_2 - y_{2i})^2 + \\ & + (a_{31}x_{1i} + a_{32}x_{2i} + a_{33}x_{3i} + t_3 - y_{3i})^2. \end{aligned} \quad (2.14)$$

Давайте определим новые координаты x_{ki} через старые координаты x_{ki} следующим образом

$$x_{ki} = x_{ki} - \frac{1}{n} \sum_{j=1}^n x_{kj}, \quad k = 1, \dots, 3, \quad i = 1, \dots, n. \quad (2.15)$$

Аналогично координаты точки y_{ki} выразим через старые координаты y_{ki} следующим образом

$$y_{ki} = y_{ki} - \frac{1}{n} \sum_{j=1}^n y_{kj}, k = 1, \dots, 3, i = 1, \dots, n. \quad (2.16)$$

Давайте определим коэффициенты $\alpha_i, \beta_i, \gamma_i, \varphi_i$ и ψ_i для $i = 1, \dots, n$:

$$\begin{aligned} \alpha_i &= x_{2i} - \frac{x_{1i}}{\sum_{j=1}^n x_{1j}^2} \sum_{j=1}^n x_{2j} x_{1j}, \\ \beta_i &= x_{3i} - \frac{x_{1i}}{\sum_{j=1}^n x_{1j}^2} \sum_{j=1}^n x_{3j} x_{1j}, \\ \gamma_i &= y_{1i} - \frac{x_{1i}}{\sum_{j=1}^n x_{1j}^2} \sum_{j=1}^n y_{1j} x_{1j}, \\ \varphi_i &= \beta_i - \frac{\alpha_i}{\sum_{j=1}^n \alpha_j^2} \sum_{j=1}^n \beta_j \alpha_j, \\ \psi_i &= \gamma_i - \frac{\alpha_i}{\sum_{j=1}^n \alpha_j^2} \sum_{j=1}^n \gamma_j \alpha_j. \end{aligned} \quad (2.17)$$

В работе сделано предположение для частного случая, вырожденное решение задачи ИСР для аффинных преобразований, когда все элементы облаков точек лежат в одной плоскости. Тогда определим элементы первой строки матрицы A , которые минимизируют функционал $J(A)$:

$$\begin{aligned} a_{11} &= \frac{\sum_{i=1}^n (y_{1i} - a_{12}x_{2i} - a_{13}x_{3i}) x_{1i}}{\sum_{i=1}^n x_{1i}^2}, \\ a_{12} &= \frac{\sum_{j=1}^n \gamma_j \alpha_j - a_{13} \sum_{j=1}^n \beta_j \alpha_j}{\sum_{j=1}^n \alpha_j^2}, \\ a_{13} &= \frac{\sum_{k=1}^n \varphi_k \psi_k}{\sum_{k=1}^n \varphi_k^2}. \end{aligned} \quad (2.18)$$

Далее в параграфе 4.2. будет показана правомочность данного предположения. Аналогичным образом получим формулы для второй и третьей строки матрицы преобразования поворот A . Для удобства введем комбинированную матрицу $\hat{M}$ аффинного преобразования, содержащую

компоненты матрицы поворота A и вектора переноса T , и рассмотрим вариационную задачу ICP с метрикой точка-плоскость. На первом шаге работы предложенного алгоритма ICP комбинированная матрица $\hat{M}$ инициализируется значениями с использованием метода RANSAC. В каждой итерации алгоритма ICP определяются соответствующие точки в облаках точек X и Y с помощью метода ближайших соседей, k -d дерева. Далее получим решение вариационной задачи ICP точка-плоскость, минимизируя значение ошибки выравнивания:

$$E(R, T) = \operatorname{argmin}_{\hat{M}} \left[\begin{aligned} &\alpha \frac{1}{W} \left(\frac{1}{|A_f|} \sum_i w_i |\hat{M} F_s^i - F_d^i|^2 \right) + \\ &+ (1 - \alpha) \frac{1}{W} \left(\frac{1}{|A_d|} \sum_j w_j \left| ((\hat{M} s_j - d_j) n_j) \right|^2 \right) \end{aligned} \right], \quad (2.19)$$

где первое слагаемое в функционале производит операцию минимизации для визуально связанных характеристик окружающей среды, а второе слагаемое производит операцию минимизации для данных глубины, w_i и w_j – веса, которые связаны с семантическими характеристиками окружающей среды, n_j – нормаль к точкам целевого облака точек [72]. Функции F_s^i и F_d^i в первом слагаемом вычисляют проекции особых точек (визуально связанных характеристик) из системы координат трехмерной сцены в систему координат RGB-D камеры. Весовой коэффициент α выбирается пользователем эмпирическим образом и определяет условный вес первого и второго слагаемых при совместном решении вариационной задачи.

2.3 Сравнительный анализ методов регистрации данных на основе ICP

Проверим полученные теоретические выводы в параграфе 2.2 в серии реальных вычислительных экспериментов с точки зрения точности решения регистрации, скорости сходимости и вычислительной сложности [32]. В диссертационном исследовании для проведения экспериментов используются:

- камеры видимого диапазона Beward B2720 (2 шт., одна камера с широким углом обзора, другая камера с узким углом обзора);
- RGB-D камера Kinect 2.0 (1 шт.);
- времяпролетная камера (ToF) MESA SR 4000 (1 шт.) для построения 3D плотных карт.

Данные мультисенсорные датчики установлены на мобильных платформах Husky UGV (1 шт.) и Odyssey 6 Robotics (2 шт.). Общая схема экспериментального комплекса представлена на рисунке 2.7.

Все реальные данные в исследовании получены для помещения – коридор Челябинского государственного университета в качестве сцены. Для исследования точности и сходимости предложенного комбинированного метода регистрации данных было проведено два эксперимента:

- в первом случае человек управлял самостоятельно мобильной платформой с установленными на ней камерой Beward B2720 и/или камерой Kinect 2.0 и двигался вдоль коридора университета;
- во втором случае камера MESA SR 4000 и/или камера Kinect 2.0 были установлены на мобильные платформы Husky UGV и/или Odyssey 6 Robotics, при этом платформы самостоятельно передвигались в коридоре при помощи метода EKF SLAM, реализация которого была взята с сайта openslam.org.

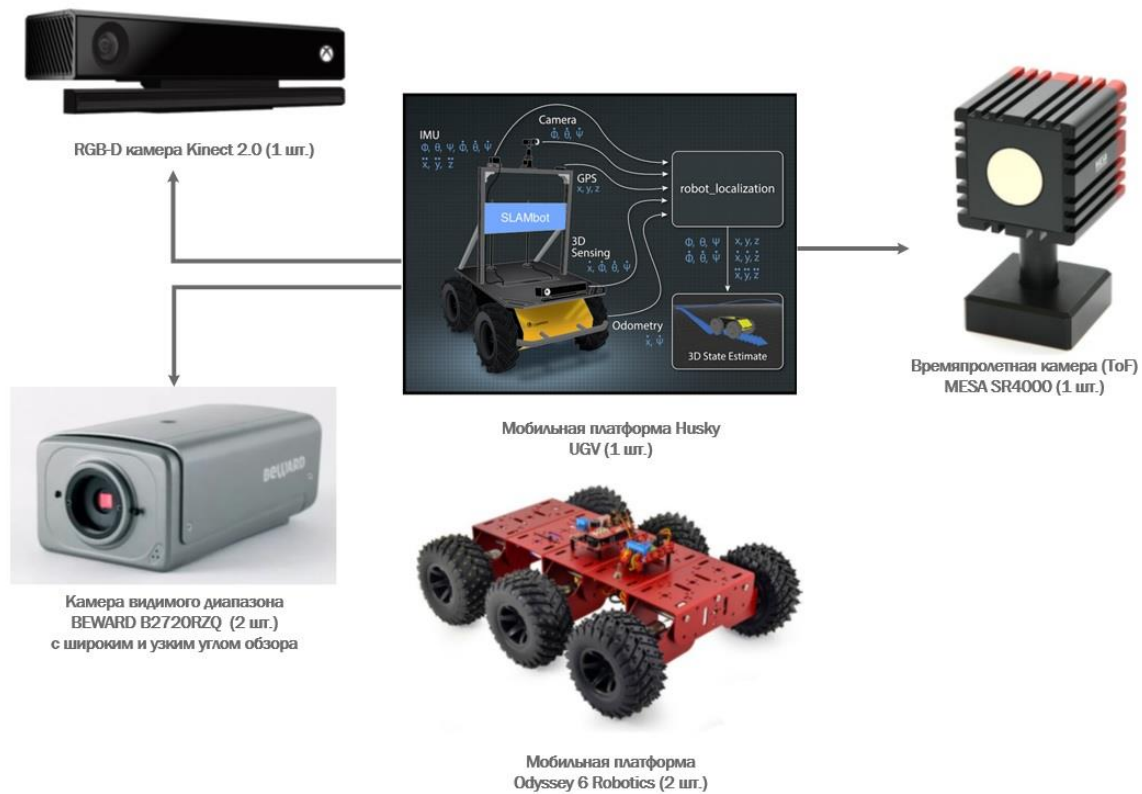


Рисунок 2.7 – Общая схема экспериментального комплекса

Эксперименты проводились в контролируемых условиях (с равномерным освещением и отсутствием шумов матричных приемников). Программное обеспечение для проведения сравнительного анализа предложенного метода и известных методов – Fast ICP и ICP точка-плоскость (метод Хорна) были взяты с сайта openslam.org. Для проведения компьютерного моделирования было захвачено с помощью камер 1650 последовательных RGB-D кадров.

В работе был проведен сравнительный анализ по точности сопоставления данных о цвете (на изображениях в кадрах) для предложенного комбинированного метода регистрации данных и для известных методов регистрации, таких как быстрый ICP (FAST ICP) и классический алгоритм ICP с метрикой точка-плоскость. Результаты компьютерного моделирования представлены в таблице 2.1, в которой каждый элемент таблицы имеет два значения: первое соответствует точности сопоставления данных (%), второе соответствует производительности (сек). На рисунке 2.8 изображены графики, которые отражают зависимость

точности методов регистрации данных от количества итераций. Точность в компьютерном моделировании измеряется с использованием среднеквадратичной ошибки, RMSE. Из рисунка 2.8 видно, что предложенный метод регистрации имеет лучшую сходимость, чем метод регистрации данных ICP с метрикой точка-плоскость или быстрый ICP.

Таблица 2.1 – Точность (в %) и производительность (сек) методов регистрации в зависимости от угла поворота вне плоскости сцены

Метод регистрации	Угол поворота (градусы)				
	5°	10°	15°	20°	25°
Быстрый ICP	91/0,06	80/0,1	73/0,11	66/0,13	56/0,15
ICP (точка-плоскость)	97/1,56	95/1,86	91/2,34	86/3,8	81/4,17
Комбинированный метод	98/1,45	99/1,78	98/2,4	96/2,95	91/3,2

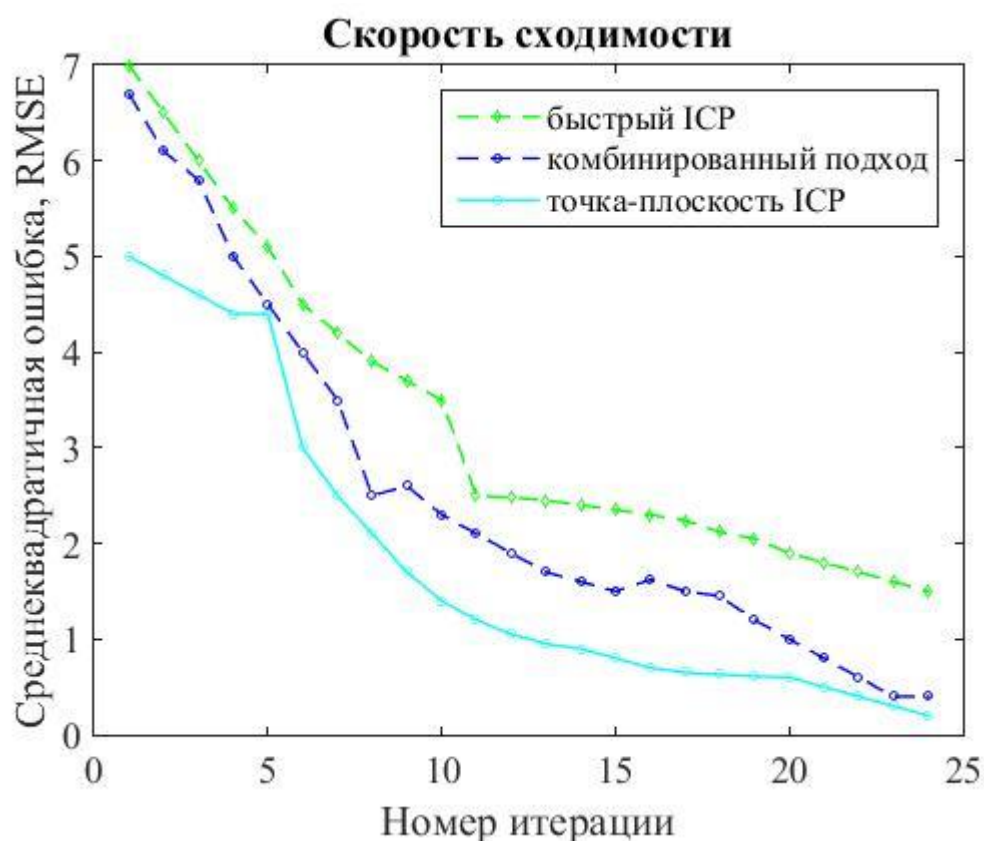


Рисунок 2.8 – Исследование сходимости методов регистрации данных на основе ICP

2.4 Методология решения проблемы SLAM на основе комбинирования визуальных и семантических характеристик неизвестной окружающей среды

В работе предложена комбинированная методология для решения проблемы SLAM, которая состоит из следующей повторяющейся последовательности шагов:

1. Исследование окружающего пространства с помощью различных мультисенсорных датчиков, получение последовательности ключевых RGB-D кадров.
2. Оценка положения мобильной платформы на сцене: определение сдвига платформы, используя сравнение данных, которые были получены на текущем шаге, с данными, полученными на предыдущих шагах.
3. Выделение из текущих RGB-D кадров признаков (особых точек).
4. Оценка карты: решение комбинированной задачи регистрации данных, сопоставление выделенных текущих признаков с признаками, которые были получены за всю хронологию наблюдений в вычислительном эксперименте.
5. Семантическая маркировка пространства на основе извлеченных особых точек (изображения) в RGB-D кадрах.
6. Уточнение карты и положения мобильной платформы: обновление положения платформы на основе визуальной и семантической информации за всю историю наблюдений.
7. Исследование полученного решения (пути) на «замыкание цикла».
8. Объединение локальных карт: обновление трехмерной карты окружающей среды, используя положения выделенных признаков и мобильной платформы за всю хронологию наблюдений в вычислительном эксперименте.

Представим графическую интерпретацию работы предложенной комбинированной методологии SLAM (Рис. 2.9).

Уровень и способ решения поставленных задач отвечает мировому. Литература, посвященная вопросам решения поставленной проблемы для

методологии Visual SLAM, содержит большое количество статей, использующих вариационные функционалы для решения задачи регистрации изображений или трехмерных облаков точек, полученных в результате сканирования окружающего пространства при помощи различных мультисенсорных датчиков. В отличие от известных методологий предлагаемая методология основана на комбинировании мультимодальных визуальных и семантических характеристик неизвестной контекстуально сложной окружающей среды и в отличие от существующих методологий использует мультисенсорную и семантическую информацию на всех ключевых этапах методологии SLAM для повышения ее точности.



Рисунок 2.9 – Графическая интерпретация комбинированной методологии SLAM

Известно, что сходимость и точность классических алгоритмов ICP, предложенных в работах Бесла П., Чена Ю., Хорна Б., может быть существенно улучшена. Предложенные методы ICP характеризуют следующие недостатки: во-первых, классические методы ICP не учитывают при анализе и обработке информации локальную форму поверхности вокруг каждой точки в трехмерном облаке, а во-вторых, известные методы регистрации данных на основе ICP

обладают большой вычислительной сложностью, при этом самая затратная операция – поиск ближайших точек. Классический метод регистрации на основе ICP решает вариационную задачу, связанную с подбором такого геометрического преобразования, которое наилучшим образом сопоставляет соответствующие точки в двух облаках точек на сцене. Геометрическое преобразование может быть аффинным или ортогональным. Выбор вида функционала в задаче оптимизации приводит к различным методам регистрации облаков точек. Наиболее используемыми являются поиск соответствия между парой облаков – точка-точка и поиск соответствия – точка-плоскость. В современных методах часто вместо соответствия точка к точке используется соответствие точка к плоскости. «При таком подходе ошибка, которую минимизирует метод, является расстоянием между точками текущего скана и касательными прямыми к точкам опорного скана. Однако подход точка-плоскость накладывает более жесткие ограничения на структуру реконструируемого пространства, такой вариант метода ICP имеет плохую сходимость на сценах с малым количеством геометрических ограничений» [13]. Для класса ортогональных преобразований решение задачи точка-точка в явном виде представлено в классических работах Хорна Б. Отметим, что методы, основанные на подходе точка-плоскость, являются более робастными к шуму датчиков, но для этих методов для класса ортогональных преобразований решение задачи в явном виде пока не найдено. Это усложняет применение методов данного класса в задачах, где требуется производить регистрацию в масштабе реального времени. Метод ICP использует два базовых шага: установление соответствия между точками двух облаков и решение вариационной задачи нахождения оптимального геометрического преобразования, совмещающего данные двух облаков точек. В рамках диссертационной работы разработаны методы, улучшающие качество работы обоих основных шагов метода. Данные о характерных признаках изображений были применены на двух этапах работы алгоритма ICP: для задания начальных значений итеративного алгоритма ближайших точек и при решении вариационной задачи (5-6 этапы алгоритма ICP, связанные с сведением к минимуму ошибки метрики L_2). Информация о

семантических свойствах объектов на сцене в предлагаемой методологии используется при решении следующих задач: регистрация данных; навигация и определение положения мобильной платформы; планирование траектории. Семантическая информация о классификации объектов на трехмерной сцене была использована для улучшения сходимости предложенного комбинированного метода регистрации: данные о семантических характеристиках были связаны со значениями весов в функционале вариационной задачи ICP. Конкретные значения весов для классов на трехмерной сцене задает пользователь системы для каждой конкретной сцены и в каждой эталонной базе данных.

В предложенной комбинированной методологии SLAM при решении вариационной задачи ICP в замкнутой форме используется функционал, позволяющий находить совместное решение на основе комбинации данных о характерных точках (визуально связанных характеристиках), данных в виде трехмерного облака точек и данных о семантических характеристиках окружающей среды.

В работе разработан алгоритм представления окрестности особых точек в виде вектора признаков (дескриптора) с высокими характеристиками с точки зрения производительности (ДЛ ГНГ, глава 3), который использует процедуру пирамидального разложения изображения и рекурсивное вычисление ГНГ. Данные о характерных точках были применены в предложенной комбинированной методологии для решения проблемы одновременной навигации и построения карты на всех ключевых шагах: при решении проблемы «замыкания цикла» и при решении задачи глобальной оптимизации; при решении вариационной задачи ICP при регистрации данных; при определении положения мобильной платформы на сцене в каждый момент времени.

Таким образом, для решения проблемы одновременной навигации и картографирования необходимо решить следующие задачи:

1. Калибровка мультисенсорных сенсоров.
2. Локальная пространственно-неоднородная обработка сигналов в спектральной области.

3. Сжатие данных и получение цифровых изображений.
4. Восстановление изображений от искажений.
5. Определение начальных весов семантических маркеров на сцене.
6. Выделение особых точек на изображениях.
7. Сопоставление особых точек на последовательности ключевых RGB-D кадров.
8. Комбинированное решение задачи регистрации данных для аффинных и ортогональных преобразований.
9. Реконструкция трехмерной сцены.
10. Идентификация семантических маркеров на сцене. Расчет весов семантических маркеров.
11. Определение положения мобильной платформы на сцене.
12. Оценка карты.
13. Решение проблемы «замыкания цикла».
14. Коррекция положения мобильной платформы на сцене.
15. Уточнение сцены.
16. Объединение локальных карт. Глобальная оптимизация.

Тогда можно представить детальную графическую интерпретацию комбинированной методологии для точного решения проблемы SLAM [11] (Рис. 2.10). Предлагаемая методология основана на динамическом сопоставлении и регистрации последовательности изображений и карт глубины и расширенной Калмановской фильтрации. В работе информация о семантических свойствах объектов предметной области (окружающей среды) используется в процессе решения задачи реконструкции трехмерной сцены, задачи навигации мобильной платформы в неизвестном пространстве и задачи планирования траектории движения мобильной платформы.

В работе предложен точный итеративный метод поиска преобразований для группы аффинных преобразований между двумя облаками точек на основе алгоритма ICP, который инвариантен как к поворотам (классический вариант), так и к произвольному масштабу (глава 4).

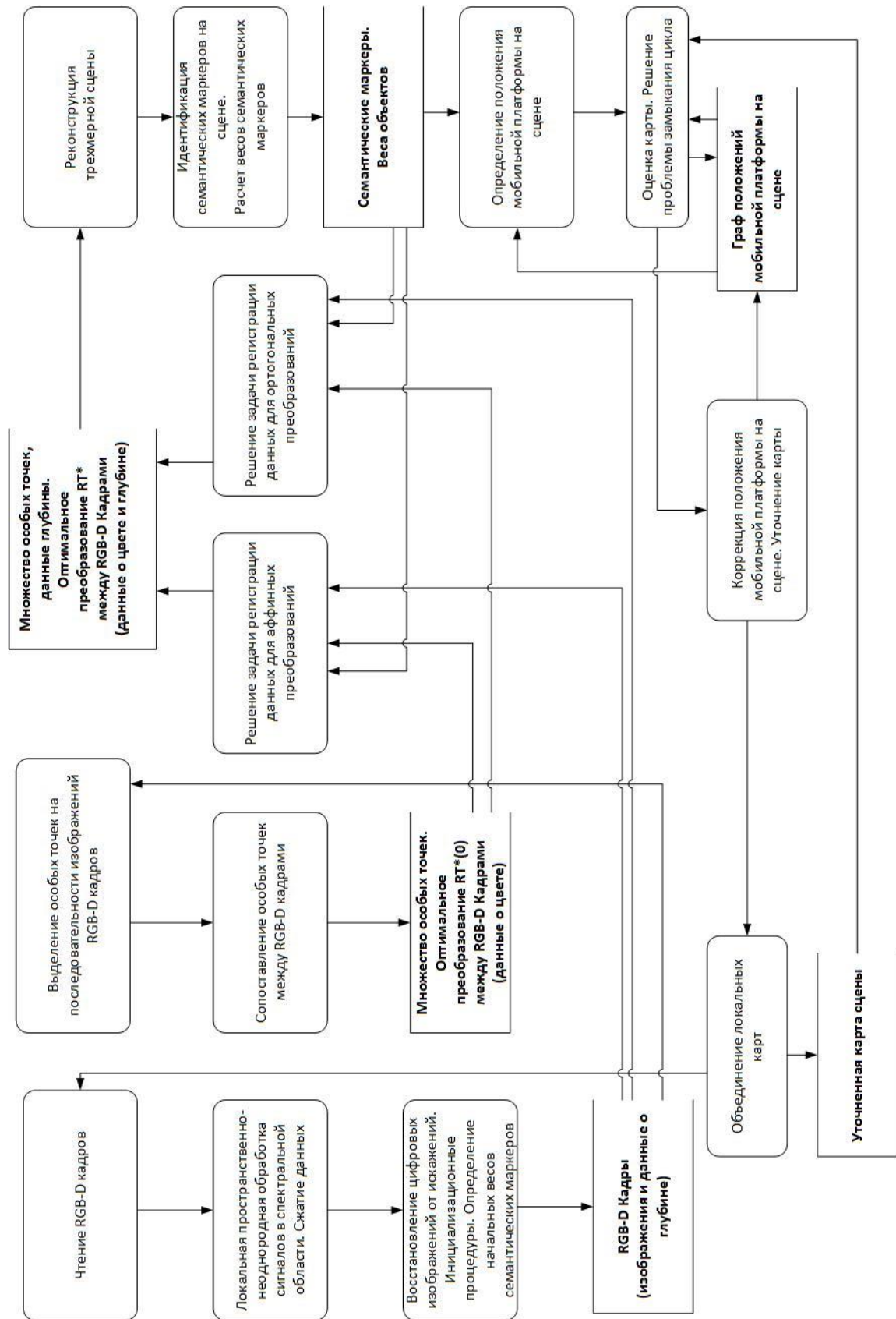


Рисунок 2.10 – Детальная графическая интерпретация комбинированной методологии SLAM

Данный метод используется для сопоставления и регистрации облаков точек с произвольным пространственным разрешением и масштабом относительно друг друга. В работе найдено точное аналитическое решение задачи точка-плоскость для группы ортогональных преобразований и на его основе разработан ригидный итеративный метод регистрации на основе алгоритма ICP, что позволяет улучшить точность и скорость регистрации трехмерных сцен (глава 4).

В работе предложен метод реконструкции трехмерной карты доступной окружающей среды и определения положения мобильной платформы в относительной системе координат на основе истории положений камеры и перемещений мобильной платформы, семантических маркеров и совмещении полученных трехмерных карт глубины, учитывающий точность их наложения и геометрические связи между различными снимками одной сцены, получены формулы для оценки вектора состояния при движении мобильной платформы по серии зашумленных измерений и семантических маркеров (глава 5). В данной работе использован подход объединения локальных карт на основе знаковой функции расстояния для непараметрического представления различных локальных карт. Для более надежного восстановления глобальной карты поверхности из зашумленных локальных карт используется норма L_1 для усеченных знаковых функций, используя постановку задачи регуляризации полной вариации.

Для решения задач сжатия изображений, получаемых с множества датчиков на мобильной платформе, был разработан быстрый рекурсивный алгоритм для вычисления дискретного косинусного преобразования (ДКП). Был использован, предложенный Кобером В.И., непараметрический метод фильтрации и восстановления изображений, искаженных аддитивным фоновым шумом, импульсными помехами и мультипликативными интерференциями, из последовательности пяти наблюдаемых изображений, полученных микросканированием камеры (глава 6). Применение методов сжатия и восстановления изображений позволяет существенно повысить точность и сходимость всей комбинированной методологии решения проблемы одновременной навигации и картографирования.

Для отслеживания повторного прохождения мобильной платформой одного и того же места на карте, на каждом шаге проводится сравнение положения и состояния мобильной платформы с предыдущими шагами, и в случае совпадения данных производится корректировка движения и уточнение трехмерной карты окружающей среды. В работе взято за основу и улучшено одно из известных решений этой проблемы на основе подхода корзина слов (Bags of Binary Words) (Рис. 2.11). Точность подхода улучшена за счет использования в процедуре определения «замыкания цикла» предложенного в работе дескриптора на основе ГНГ, вычисляемых в скользящем окне (глава 7).

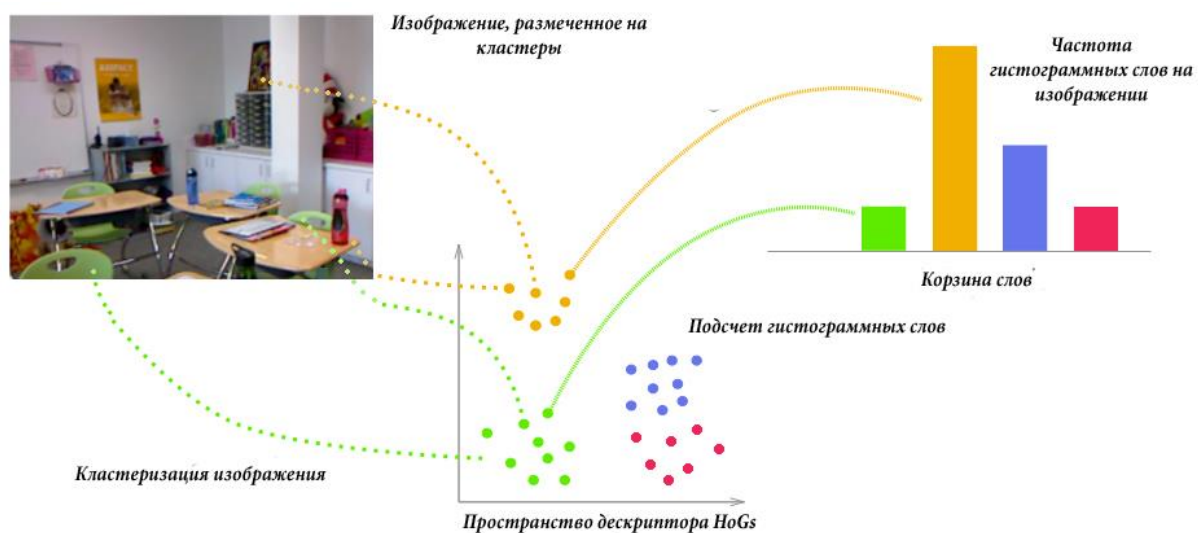


Рисунок 2.11 – Графическая интерпретация подхода корзина слов

В работе рассмотрен эффективный эвристический алгоритм планирования траектории мобильной платформы (глава 5) в неизвестном пространстве на основе навигации по семантическим маркерам и метрико-топологического графа, который использует информацию о семантических свойствах трехмерной сцены для поиска локально-оптимальной траектории на сцене. Как будет показано в главе 5, использование информации о семантике объектов на сцене приводит к сокращению вычислительной и алгоритмической сложности метода планирования траектории мобильной платформы. В главе 8 будет рассмотрено применение данного алгоритма планирования траектории в различных практических

приложениях предлагаемой комбинированной методологии решения проблемы одновременной навигации и построения карты.

Комбинированная методология решения проблемы одновременной навигации и построения карты имеет следующие преимущества:

1. 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 характерных точках.

2. Информация о семантических характеристиках окружающей среды извлекается из данных о цвете (изображений в последовательности кадров) 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.

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

4. 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.

5. 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.

2.5 Выводы по главе

1. Предлагаемая в работе методология решения проблемы SLAM для контекстуально сложных крупномасштабных сцен основана на совместном использовании мультисенсорной динамической информации и семантической информации о объектах на сцене, пространственном совмещении последовательных кадров данных, обнаружении «замыканий цикла» (возврат мобильной платформы в ранее посещаемое ей место), выравнивании полной последовательности данных (сопоставление и регистрация последовательностей изображений и трехмерных облаков точек). Реализация задач диссертационного исследования выполнена посредством прорывных научных исследований и разработок мирового уровня в области информационных технологий. Предложенные методы и алгоритмы позволяют создать прототип системы одновременной навигации и построения карты лучшего качества, чем известные существующие системы на основе визуального SLAM, работающие в аналогичных условиях.

2. Комбинирования методология решения проблемы SLAM соответствует современному уровню развития аналогичных решений для задач автономной навигации интеллектуальных систем в неизвестном пространстве. Ключевым этапом предлагаемой методологии является решение задачи реконструкции трехмерной карты окружающей среды. В предложенной методологии SLAM при решении вариационной задачи ICP в замкнутой форме используется функционал, позволяющий находить совместное решение на основе комбинации данных об особых точках, данных в виде трехмерного облака точек и данных в виде семантических характеристик окружающей среды.

3. В работе рассмотрен частный случай решения комбинированной вариационной задачи ICP, когда данные в трехмерных облаках точек находятся в одной плоскости сцены. На основе компьютерного моделирования и вычислительного эксперимента был сделан вывод, что предложенный

комбинированный метод регистрации имеет лучшую сходимость, чем известные методы регистрации данных, при этом предложенный метод показывает близкие по точности результаты с методом регистрации на основе алгоритма ICP, который использует метрику точка-плоскость. Следовательно, предложенный метод решения комбинированной вариационной задачи ICP является перспективным для решения проблемы SLAM для контекстуально сложных крупномасштабных сцен. Далее в работе будет найдено решение данной вариационной задачи ICP в замкнутой форме для различных классов геометрических преобразований.

4. Предложенная методология в отличие от существующих методологий использует мультисенсорную и семантическую информацию на всех этапах решения проблемы SLAM для повышения точности получаемых результатов. Разработанная методология позволяет осуществлять реконструкцию трехмерной карты неизвестной окружающей среды в замкнутой форме с использованием семантических характеристик объектов на сцене и доступной мультисенсорной информации, поступающей от датчиков при произвольном изменении их положения в трехмерном пространстве в контролируемых и неконтролируемых условиях.

5. Предложенная в исследовании комбинированная методология представляет сложную научно-техническую проблему, поэтому на отдельных этапах исследования будут использоваться либо известные методы и подходы, либо известные научные результаты, полученные в рамках научно-исследовательских работ в ходе выполнения грантов, где автор работы являлся руководителем или основным исполнителем. При выполнении диссертационного исследования были использованы различные научные результаты, методы и алгоритмы, полученные в результате выполнения грантов РФФИ № 16-08-00342 (руководитель Вохминцев А.В.), РНФ № 15-19-10010 (руководитель Кобер В.И.), РФФИ № 18-37-20032 (руководитель Яковлев К.С.), Государственное задание (проектная часть) Министерство образования и науки РФ 2.1766.2014/К (руководитель Кобер В.И.), Государственное задание (проектная часть) Министерство образования и науки РФ 2.1743.2017.ПЧ (руководитель Кобер В.И.).

Отметим ряд наиболее важных известных методов и подходов, которые были использованы в исследовании:

- итеративный алгоритм ближайших точек (Бесл П., Чен Ю., Хорн Б.);
- методология EKF SLAM (Транс С., Монтемерло М., Нейра Дж., Тардос Дж. Д.);
- подход к оптимизации графа положений мобильной платформы и обнаружению «замыканий цикла» на основе метода корзина слов (Нистер Д., Климент Л.);
- итерационные методы восстановления и обнаружения объектов на восстановленном изображении на основе оптимальной согласованной фильтрации (Кобер В.И., Маковецкий А.Ю., Пратт У., Диас-Рамирес В.);
- непараметрический метод фильтрации и восстановления изображений, искаженных аддитивным фоновым шумом, импульсными помехами и мультипликативными интерференциями с использованием микросканирования камеры (Кобер В.И., Диас-Рамирес В.);
- метод планирования траектории на основе метрико-топологического графа и алгоритма HGA* (Яковлев К.С.);
- методы семантической маркировки пространства на основе запутанных трехмерных лесов (3DEF) и детектора объектов YOLO (Антонелло М., Вольф Д., Редмон Д.);
- метод локализации объектов на изображении на основе ГНГ (Кобер В.И., Диас-Рамирес В.);
- методы для локально-адаптивной обработки изображений на основе скользящих ДКП и ДСП (Кобер В.И.).

6. Основные положения, выносимые на защиту (п.1, п. 2, п.3, п.4 и п.6), являются личным вкладом диссертанта в исследование и опубликованные работы. Результат п.5 основных положений, выносимых на защиту, получен автором совместно с Кобером В.И., Мельниковым А.В. и является неделимым.

7. Остальные научные результаты получены автором диссертационной работы совместно с Мельниковым А.В, Соченковым И.В., Козько А.А, Тимченко М.С. и другими соавторами.

8. Основные результаты по данной главе опубликованы в работах [11, 20, 32].

ГЛАВА 3 МЕТОДЫ ПОИСКА И СОПОСТАВЛЕНИЯ ОСОБЫХ ТОЧЕК НА ИЗОБРАЖЕНИЯХ НА ОСНОВЕ ДВУМЕРНЫХ ДЕСКРИПТОРОВ

На сегодняшний день предложено множество методов распознавания изображений и соответствующих им методов сопоставления данных, среди которых следует наиболее часто использоваться следующие методы:

1. Scale Invariant Feature Transform (SIFT) [95] – масштабно-инвариантное преобразование признаков (особые точки).
2. Speeded-Up Robust Features (SURF) [96] – ускоренные устойчивые признаки.
3. Oriented FAST and Rotated BRIEF (ORB) [97] является комбинацией модифицированных дескрипторов FAST и BRIEF. Первый дескриптор извлекает особые точки из изображения, второй дескриптор определяет свойства особых точек в формате бинарной строки.

Данные дескрипторы извлекают характерные признаки из изображений, производят сопоставление данных и имеют высокую алгоритмическую сложность, поэтому их нельзя использовать для решения поставленной задачи – одновременной навигации и построения карты для крупномасштабных контекстуально сложных сцен, где с одной стороны обрабатываются большие потоки данных, а с другой стороны требуется получать результаты в реальном масштабе времени. В данной главе предложен метод представления окрестности особых точек в виде вектора признаков (дескриптора) с высокими характеристиками с точки зрения вычислительной сложности алгоритма на основе вычисления гистограмм направленных градиентов, который применим для изображений, имеющих геометрические искажения (поворот, перенос, масштаб). Разработан алгоритм поиска особых точек на изображениях последовательности кадров на основе дескриптора локальной гистограммы ориентированных градиентов с рекурсивной процедурой вычисления.

3.1 Алгоритм извлечения особых точек на основе дескриптора с рекурсивным вычислением гистограмм направленных градиентов

При цифровой обработке изображений для сопоставления данных часто применяют методы на основе гистограмм ориентированных градиентов (ГНГ). ГНГ позволяют проводить сопоставление данных с высокой различительной способностью. Давайте обозначим через $\{x_i, y_i\}$ – множество координат, соответствующих центрам круглых дисков с радиусами r_i . В литературе по цифровой обработке изображений данные круглые диски называют скользящим окном. Экспериментальным путем установлено, что для достижения высокой точности процедуры сопоставления изображений необходимо использовать от 2 до 4 круглых скользящих окон. В диссертационном исследовании разработан алгоритм представления окрестности особых точек в виде вектора признаков на основе гистограмм направленных градиентов с рекурсивной процедурой вычисления, апробированный для решения проблемы SLAM [10]. Предложенный в работе алгоритм извлечения особых точек основан на двумерном дескрипторе, который использует несколько скользящих окон и рекурсивную процедуру ГНГ [174, 175]. Для построения ГНГ необходимо определить значение градиентов для каждого пикселя на основе операторов формы:

$$gx = [-1 \quad 0 \quad 1], \quad gy = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}. \quad (3.1)$$

Далее, зная значение градиентов, можно вычислить значение магнитуды для каждого пикселя изображения:

$$mag(x, y) = \sqrt{gx^2 + gy^2}. \quad (3.2)$$

Аналогично на основе значений градиентов можно вычислить значение ориентации для каждого пикселя изображения:

$$ori(x, y) = \arctan\left(\frac{gy}{gx}\right). \quad (3.3)$$

Значение ориентации в формуле (3.3) изменяется в диапазоне $[0^\circ, 360^\circ]$. Давайте обозначим набор контейнеров Q в соответствии со стандартным отклонением значения шума σn . Тогда ориентация в каждом пикселе будет квантована следующим образом

$$\varphi(x, y) = \left\lfloor \frac{Q}{360} ori(x, y) + \frac{1}{2} \right\rfloor, \quad (3.4)$$

где фактор « $1/2$ » позволяет разместить данные о величине угла из начальной и конечной области ГНГ в первый контейнер. В выражении (3.4) выполняется условие циклической ориентации: ошибка ориентации градиента может установить значения как ближе к 0° , так и наоборот, к 360° .

Пусть $\{W_i, i = 1, \dots, M\}$ – набор медианных фильтров (МФ) на основе круглых скользящих окон на эталонном изображении следующего вида

$$W_i = \{(x, y) \in R^2: (x - x_i)^2 + (y - y_i)^2 \leq r_i^2\}, \quad (3.5)$$

где (x_i, y_i) – координаты центра МФ и r_i – радиус i -го диска, соответствующего МФ. Данный набор круглых дисков образует на плоскости некоторую геометрическую фигуру с определенной топологией (Рис. 3.1). Далее вычислим относительные расстояния и углы между центрами круглых скользящих дисков для данной топологии. В предложенном двумерном дескрипторе вычислим ГНГ в круглых областях и затем используем их для сопоставления изображений. Для каждого положения i -го круглого окна из набора на анализируемом изображении определим значение градиентов на основе оператора Собеля. Определим ГНГ в каждом положении i -го круглого скользящего окна на изображении

$$HOG_i(\alpha) = \begin{cases} \sum_{(x,y) \in W_i} \delta(\alpha - \varphi_i(x, y)), & \text{if } Mag_i(x, y) \geq Med \\ 0, & \text{otherwise} \end{cases}, \quad (3.6)$$

где $\alpha = \{0, \dots, Q - 1\}$ – порядковый номер дискретного уровня ГНГ, используется для квантования значений углов по различным направлениям в предлагаемом дескрипторе; $\{Mag_i(x, y): (x, y) \in W_i\}$ – значения модулей градиентов; $\{\varphi_i(x, y): (x, y) \in W_i\}$ – значение угла направлений; Med – медиана

модулей градиентов круглых МФ; $\delta(z) = \{1, \text{если } z = 0; 0, \text{ в противном случае}\}$ и представляет из себя символ Кронекера.



Рисунок 3.1 – Пример набора круглых окон для эталонного изображения

Вычисление ГНГ в каждом положении i -го круглого скользящего окна в формуле (3.6) представляет из себя процедуру с высокой алгоритмической сложностью: (πr_i^2 – количество вычислительных операций суммирования элементов). Для преодоления данного недостатка при определении значений ГНГ используется рекурсивная процедура решения уравнения (3.6). В этом случае алгоритмическая сложность при вычислении значений ГНГ может быть оценена как $2\pi r_i$ вычислительных операций суммирования элементов

$$HOG_i^k(\alpha) = HOG_i^{k-1}(\alpha) - \sum_{(x,y) \in OutP_i^{k-1}} \delta(\alpha - Out_{\phi_i}^{k-1}(x,y)) + \sum_{(x,y) \in InP_i^k} \delta(\alpha - Int_{\phi_i}^k(x,y)). \quad (3.7)$$

Для обеспечения инвариантности предложенного двумерного дескриптора к преобразованию «поворот» в исследовании используется нормализованный корреляционный оператор, при помощи которого осуществляется процедура сравнения объектов на анализируемой двумерной сцене с соответствующей ей ГНГ эталонного изображения. Значения нормализованной центрированной ГНГ эталонного изображения определяются по формуле

$$\overline{HOG_i^R}(\alpha) = \frac{HOG_i^R(\alpha) - Mean^R}{\sqrt{Var^R}}, \quad (3.8)$$

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

$$C_i^k(\alpha) = IFT \left[\frac{HS_i^k(\omega) HR_i^*(\omega)}{\sqrt{Q \sum_{q=0}^{Q-1} (HOG_i^k(q))^2 - (HS_i^k(0))^2}} \right], \quad (3.9)$$

где $HS_i^k(\omega)$ – преобразование Фурье ГНГ внутри i -го окна входной сцены, $HR_i(\omega)$ – преобразование Фурье $\overline{HOG_i^R}(\alpha)$, знак $*$ – комплексное сопряжение.

В исследовании мерой подобия двух различных ГНГ является корреляционный пик, который может быть получен по следующей формуле

$$P_i^k = \max_{\alpha} \{C_i^k(\alpha)\}. \quad (3.10)$$

Заметим, что значения корреляционных пиков в формуле (3.10) изменяются в интервале $[-1, 1]$. Для уменьшения вычислительной сложности в данной работе значения нормализованных центрированных ГНГ и вычисление преобразований Фурье вычисляются перед началом итеративной процедуры сопоставления данных.

3.2 Алгоритм быстрого сопоставления изображений

На основе математической модели двумерного дескриптора ДЛ ГНГ, представленной в параграфе 3.1, был разработан алгоритм сопоставления изображений [10, 11]. На рисунке 3.2 представлена блок-схема данного алгоритма. Давайте дадим словесное описание данного алгоритма.

Шаг 1. Давайте определим набор круглых скользящих окон (КСО) $\{W_i, i = 1, \dots, M\}$ для заданного эталонного изображения. При этом отметим, что все скользящие окна имеют одинаковый постоянный радиус r . Далее определим расстояние между центрами всех КСО на изображении по следующей формуле:

$$D_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}, i = 1, \dots, M; j = i + 1, \dots, M. \quad (3.11)$$

Шаг 2. Далее выполним операцию сжатия для ГНГ $\{\overline{HoG}_i^R(\alpha), i = 1, \dots, M\}$. Определим число уровней квантования ГНГ Q . Выполним предварительную обработку данных: вычислим преобразования Фурье $\{CHR_i(\omega), i = 1, \dots, M\}$. Далее выполним процедуру построения треугольной сетки на анализируемом изображении и вычислим в каждом треугольнике соответствующие углы (вершинами данных треугольников являются центры КСО):

$$\gamma_i = \cos^{-1} \left[\frac{D_{i,j+1}^2 + D_{i,j+2}^2 - D_{i,j+3}^2}{2D_{i,j+1}D_{i,j+2}} \right], i = 1, \dots, M - 2. \quad (3.12)$$

Шаг 3. Далее получим «уменьшенное» эталонное изображение с шагом равным L . Для получения «уменьшенного изображения» применим процедуру down-sampling вдоль всех столбцов и строк изображения. На выходе процедуры down-sampling будет получено L^2 «уменьшенных» изображений. Далее выполним предварительную обработку данных: для каждого i -го КСО эталонного изображения вычислим значения нормализованных центрированных ГНГ (НЦ ГНГ) $\{\overline{HoG}_i^R(\alpha), i = 1, \dots, M\}$ по формуле (3.8) и вычислим соответствующее преобразование Фурье $\{HR_i(\omega), i = 1, \dots, M\}$.

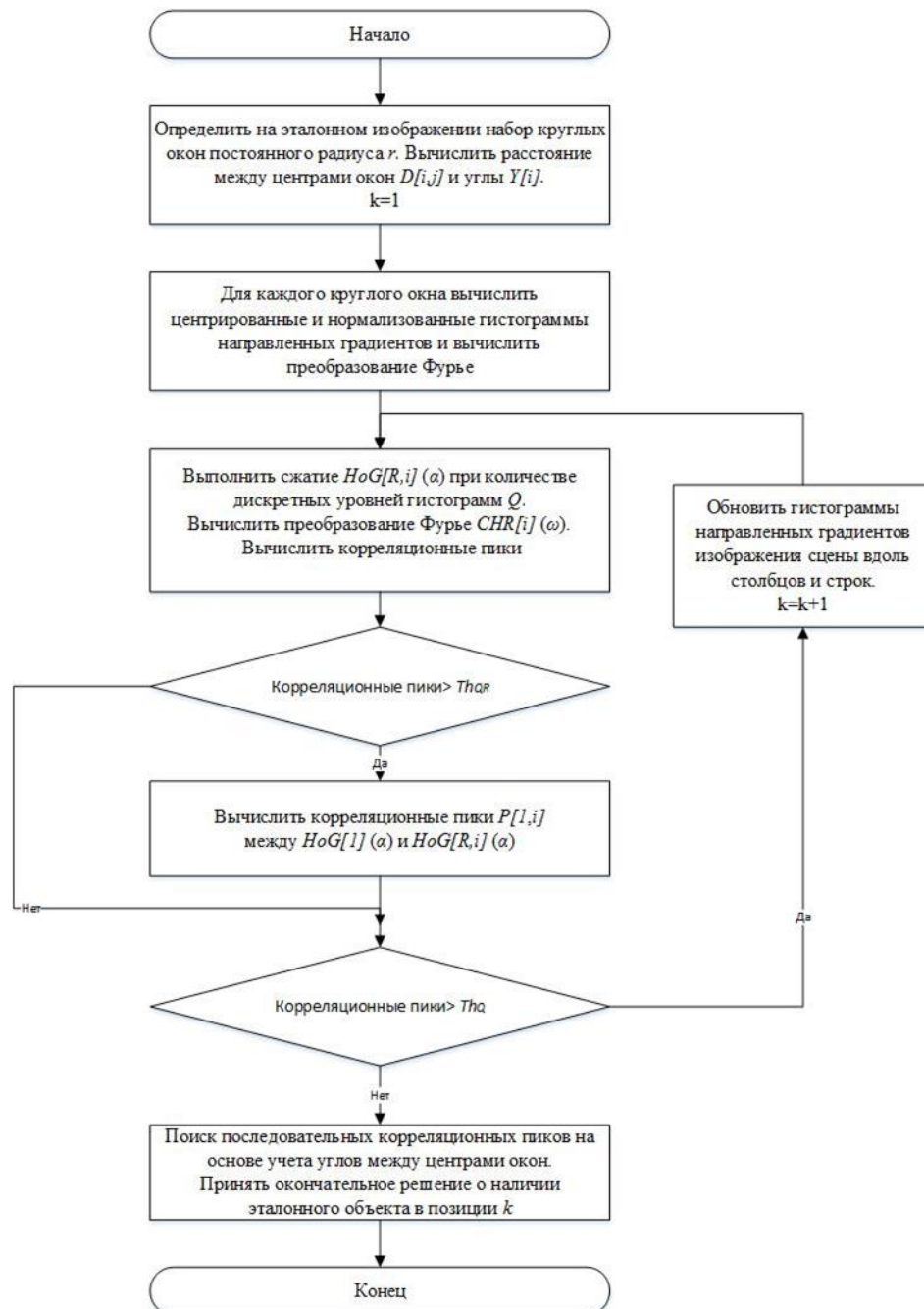


Рисунок 3.2 – Блок-схема алгоритма быстрого сопоставления изображений

Шаг 4. Далее выполним процедуру сжатия для ГНГ $\{\overline{HoG}_i^R(\alpha), i = 1, \dots, M\}$. Затем по формуле (3.8) вычислим преобразование Фурье $\{CHR_i(\omega), i = 1, \dots, M\}$ для количества уровней ГНГ равного Q.

Шаг 5. Далее получим «уменьшенное» изображение сцены с шагом, равным L (1-L). Далее аналогично Шагу 3 применим процедуру down-sampling и в результате на выходе получим L2 «уменьшенных» изображений. Положим k=1.

Шаг 5.1 Выполнить процедуру сопоставления данных для всех КСО в соответствующем изображении.

Шаг 5.2 Для 1-L определим значение его ГНГ $HoG^1(\alpha)$. Далее вычислим прямое преобразование Фурье $HS^1(\omega)$ для этого изображения.

Шаг 5.3 Выполним процедуру сжатия для ГНГ $HoG^1(\alpha)$ при числе контейнеров Q, задающих уровень квантования ГНГ. Вычислим соответствующее прямое преобразование Фурье $CHS^1(\omega)$ и определим значение корреляционной функции по формуле (3.9) $CHoG^1(\alpha)$ для Q_R уровней квантования ГНГ.

Шаг 6. Вычислим корреляционные пики на основе формул (3.9) и (3.10) между $CHoG^1(\alpha)$ и $\{\overline{CHoG_i^R}(\alpha), i = 1, \dots, M\}$.

Шаг 7. Если любой из полученных на Шаге 6 корреляционных пиков больше, чем значение порога $ThQR$, задаваемого пользователем, то вычислим значения следующих корреляционных пиков $\{P_i^1, i = 1, \dots, M\}$ между $HoG^1(\alpha)$ и $\{\overline{HoG_i^R}(\alpha), i = 1, \dots, M\}$.

Шаг 8. Проверка условия для любого корреляционного пика $\{P_i^1, i = 1, \dots, M\}$: если хотя бы один пик будет больше, чем значение порога ThQ , задаваемого пользователем, то продолжаем процесс сопоставления изображений. Далее увеличиваем значение позиции k на 1. Затем запускаем процедуру обновления ГНГ изображения. Далее возвращаемся назад и повторяем Шаг 5 и Шаг 6 данного алгоритма.

Шаг 9. На основе полученных результатов примем решение о наличии «эталонного объекта» в k-ой позиции исследуемого изображения.

3.3 Компьютерное моделирование

Проведем сравнительный анализ предложенного алгоритма сопоставления изображений с точки зрения количества правильных сопоставлений в процентах (%) и вычислительной сложности в сравнении с известными методами сопоставления изображений в сек. «Для проведения тестов использовались следующие базы данных: Labeled Faces in the Wild, 3D Mask Attack Dataset, Texas 3D Face Recognition Database. Известные алгоритмы сопоставления данных были протестированы в серии различных экспериментов:

- преобразование изменения масштаба (Табл. 3.1);
- преобразование «поворот в плоскости сцены» (Табл. 3.2);
- преобразование «поворот изображения вне плоскости сцены» (Табл. 3.3)» [10, 40].

Таблица 3.1 – Точность сопоставления при использовании выбранных алгоритмов в зависимости от небольших изменений масштаба изображения сцены

Алгоритм сопоставления изображений	Значение коэффициента масштабирования				
	0.8 X	0.9 X	1.0 X	1.1 X	1.2 X
SIFT	92	95	100	98	91
SURF	79	90	99	97	92
ORB	78	79	90	83	89
Предложенный алгоритм	84	94	100	99	91

«Результаты компьютерного моделирования показали, что предложенный алгоритм превосходит известные алгоритмы сопоставления данных при выполнении «поворота изображения вне плоскости сцены», данный результат имеет большое прикладное значение для поставленной задачи – реконструкции

крупномасштабных контекстуально сложных трехмерных сцен в реальном времени» [13].

Таблица 3.2 – Точность сопоставления при использовании выбранных алгоритмов в зависимости от угла поворота изображения в плоскости изображения сцены

Алгоритм сопоставления изображений	Значение угла поворота, град.					
	5	10	15	20	25	30
SIFT	98	91	78	64	58	47
SURF	82	77	64	55	38	29
ORB	96	84	77	61	58	54
Предложенный алгоритм	83	78	74	72	76	72

Таблица 3.3 – Точность сопоставления при использовании выбранных алгоритмов в зависимости от угла поворота изображения вне плоскости изображения сцены

Алгоритм сопоставления изображений	Значение угла поворота, град.					
	45	90	135	180	225	270
SIFT	100	98	99	98	97	95
SURF	74	69	74	69	74	69
ORB	87	85	83	86	85	87
Предложенный алгоритм	100	98	96	98	97	95

Предложенный метод сопоставления изображений при вычислении ГНГ применяет набор скользящих окон из трех круглых дисков радиуса r , значение которого зависит от размера анализируемого объекта на трехмерной сцене (Рис. 3.3).

Дополнительно в результате компьютерного моделирования было показано, что производительность предложенного алгоритма пропорциональна размеру анализируемого фрагмента. Поэтому, становится возможным создание на основе

данного дескриптора методов локализации объекта (мобильной платформы) в неизвестном пространстве и методов слежения за объектами.

Приведем результаты тестов по производительности для известных алгоритмов сопоставления изображений и предложенного алгоритма: алгоритм на основе ДЛ ГНГ – 1,7 сек; на основе ORB – около 1,7 сек; на основе SIFT – 9,82 сек; на основе SURF – 0,95 сек. Таким образом, в работе предложен точный метод представления окрестности особых точек в виде вектора признаков (дескриптора) с высокими характеристиками с точки зрения производительности.

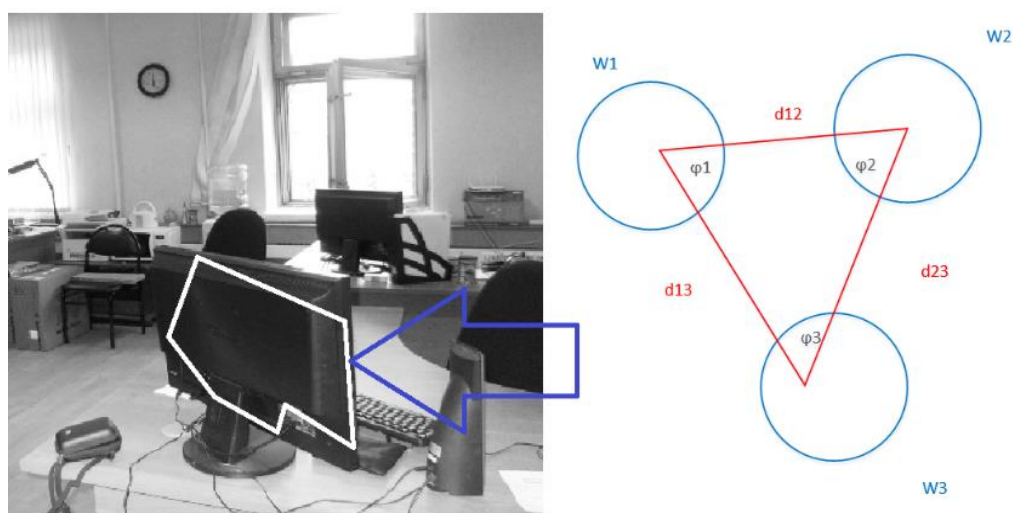


Рисунок 3.3 – Пример расположения круглых скользящих окон для анализируемого изображения (объекта в помещении)

3.4 Выводы по главе

1. В работе предложен новый алгоритм для сопоставления изображений в RGB-D кадре на основе ориентированных градиентов, который основан на рекурсивной процедуре определения значений ГНГ по трем круглым скользящим окнам и дискретизации изображения в виде пирамидального разложения на несколько связанных квантованных уровней.

2. В результате проведенных исследований и компьютерного моделирования было показано, что предложенный алгоритм сопоставления изображений на основе двумерного дескриптора ДЛ ГНГ обладает лучшими характеристиками с точки зрения точности, чем известные алгоритмы сопоставления изображений при преобразовании «поворот изображения вне плоскости сцены». Алгоритм был успешно апробирован для сцен различного класса, в том числе для трехмерных сцен с неконтролируемыми условиями (неравномерным освещением).

3. Предложенный алгоритм является инвариантным к поворотам и небольшому изменению масштаба, поэтому может успешно применяться при решении проблемы одновременной навигации и картографирования и в других приложениях робототехники, где требуется выполнять быстрое сопоставление и регистрацию данных с геометрическими искажениями в реальном масштабе времени.

4. Основные результаты по данной главе опубликованы в работах [10, 11, 40].

ГЛАВА 4 РЕКОНСТРУКЦИЯ ТРЕХМЕРНЫХ СЦЕН НА ОСНОВЕ ТОЧНЫХ РЕШЕНИЙ ВАРИАЦИОННОЙ ЗАДАЧИ РЕГИСТРАЦИИ МЕТОДА ICP

В данной главе рассматриваются комбинированные методы для реконструкции контекстуально сложных крупномасштабных трехмерных сцен для группы аффинных и для группы ортогональных преобразований на основе решения вариационной задачи ICP 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 ICP: определение соответствующих точек между парой трехмерных облаков точек и решение вариационной задачи.

Рассмотрим современное состояние проблемы. Существующие методы для решения задачи трехмерной регистрации данных принято разделять на два класса [92-94].

Методы первого класса базируются на итеративном алгоритме ближайших точек (Iterative Closest Point, ICP) [71] и используют для реконструкции трехмерной сцены инкрементальную вычислительную процедуру, связанную с решением задачи минимизации функции ошибки [75]. Недостатком методов первого класса является необходимость определения пользователем параметров инициализации для вычисления соответствующего преобразования на первом шаге работы итеративного алгоритма.

Методы второго класса используют для реконструкции трехмерной сцены различные двумерные проекции сцены и методы сопоставления особых точек с использованием рассмотренных в главе 3 дескрипторов [95-97]. Основными недостатками методов второго класса являются большая вычислительная сложность и требование наличия большого количества общих особых точек в сопоставляемых кадрах для точной реконструкции сцены.

На ключевом шаге итеративного алгоритма ближайших точек решается задача нахождения такого геометрического преобразования между парой облаков трехмерных точек (ортогонального или аффинного), которое позволяет совместить наилучшим образом эти облака точек на основе некоторой метрики близости. Данный этап метода ICP часто называют вариационной задачей алгоритма ICP. Заметим, что точность реконструкции трехмерной сцены существенно зависит от способа решения вариационной задачи и выбора метрики для оценки соответствующего геометрического преобразования.

Точность и сходимость итерационных методов при решении вариационной задачи, связанной с минимизации функции ошибки, зависит от выбора начальных значений параметров метода ICP. Как известно, итеративный алгоритм ближайших точек инициализируется начальными значениями параметров геометрического преобразования, такими как поворот, параллельный перенос и изменение масштаба. В зависимости от выбора параметров инициализации метода ICP итерационный процесс может сходиться, сходиться к локальному минимуму, сходиться с большим количеством итераций, либо не сходиться в принципе. В данной работе будет предложено решение вариационной задачи ICP в замкнутой форме, которое позволяет управлять процессом сходимости итеративного алгоритма. Классический алгоритм ICP, предложенный Беслом П. использует метрику точка-точка при минимизации функционала и может решаться с использованием метода оптимизации Левенберга-Марквардта. Известны работы [103, 105], которые рассматривают группу аффинных преобразований и развивают классический метод ICP на частные случаи, в которых трехмерное или двумерное облако точек вырожденное. Хорн Б. предложил в своих работах решения вариационной задачи ICP в замкнутой форме для группы ортогональных преобразований с использованием кватернионов [100] и ортогональных матриц [101]. Кватернионы представляют из себя гиперкомплексные числа специального вида и являются удобным математическим аппаратом для обозначения операций положения и геометрического преобразования (поворот). На основе результатов, полученных в работах Хорна Б., был определен итеративный алгоритм ближайших

точек, использующий метрику точка-точка [71]. В работе Чена Ю. [74] показано, что во многих приложениях, особенно при обработке больших объемов данных, применение метрики точка-плоскость при решении вариационной задачи ICP позволяет добиться как лучшей сходимости, так и более высокой точности при сопоставлении облаков точек. Аналогичным образом для группы ортогональных преобразований при решении вариационной задачи точка-плоскость при минимизации функционала используют метод оптимизации Левенберга-Марквардта или метод условной линеаризации углов [73].

Недавно в работе Ван дер Санде К. [107] было предложено точное решение вариационной задачи ICP для группы аффинных преобразований с использованием метрики точка-плоскость. Отметим, что решение вариационной задачи ICP для группы аффинных преобразований может использоваться как для трехмерных сцен с ригидными объектами, так и для трехмерных сцен с не ригидными объектами. Важно учитывать вид истинного геометрического преобразования (аффинного или ортогонального) между парой трехмерных облаков точек: если преобразование изначально является ортогональным, то решение вариационной задачи ICP для группы аффинных преобразований в большинстве случаев даст неверный результат. Верный результат может быть получен, если изначально соответствие между трехмерными облаками точек является близким к идеальному.

В большинстве реальных приложений робототехники пространственное отношение между облаками точками неидеальное, например, могут существовать случаи, когда большинство точек одного трехмерного облака точек соответствуют только одной точке другого трехмерного облака точек, либо соответствуют некоторому локальному подмножеству точек другого трехмерного облака точек [176]. Отметим, что аналогичное утверждение о требовании идеального отношения между трехмерными облаками точек для случая с истинным аффинным преобразованием не верно.

Решение данной условной вариационной задачи для метрики точка-точка является известным методом Хорна [100, 101]. Известны точные решения для

группы аффинных преобразований, которые используют метрику точка-плоскость при решении вариационной задачи [74].

В данной главе проведен сравнительный анализ методов регистрации относительно облака трехмерных точек на основе вариантов метода ICP, получены комбинированные решения вариационной задачи точка-точка и точка-плоскость в замкнутой форме и проведено компьютерное моделирование. В главе предложены алгоритмы для решения комбинированной вариационной задачи метода ICP для аффинных и ортогональных преобразований. Предложенные алгоритмы основаны на подобию групп геометрических элементов в облаке точек [177, 178] и комбинированном решении вариационной задачи.

4.1 Точное решение вариационной задачи ИСР для аффинных преобразований

Введем следующие определения: $X = \{x_1, \dots, x_n\}$ – исходное трехмерное облако точек и $Y = \{y_1, \dots, y_m\}$ – целевое трехмерное облако точек в $\mathbb{R}^3$. В параграфе 2.2 было показано, что итеративный алгоритм ближайших точек (ИСТ) [71-73] может быть представлен как

$$Rx_i + T, \quad (4.1)$$

где R – матрица компонентов преобразования поворот, T – вектор компонентов преобразования перенос, $i = 1, \dots, n$.

Представим функционал $J(R, T)$ в виде функции [13]

$$J(R, T) = \sum_{i=1}^n \| Rx_i + T - y_i \|^2. \quad (4.2)$$

Определим вариационную задачу ИСТ как

$$\underset{R, T}{\operatorname{argmin}} J(R, T), \quad (4.3)$$

где

$$R = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}, T = \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix}, x_i = \begin{pmatrix} x_{1i} \\ x_{2i} \\ x_{3i} \end{pmatrix}, y_i = \begin{pmatrix} y_{1i} \\ y_{2i} \\ y_{3i} \end{pmatrix}. \quad (4.4)$$

Далее заметим, что

$$\begin{aligned} J(R, T) = \sum_{i=1}^n & (r_{11}x_{1i} + r_{12}x_{2i} + r_{13}x_{3i} + t_1 - y_{1i})^2 + \\ & + (r_{21}x_{1i} + r_{22}x_{2i} + r_{23}x_{3i} + t_2 - y_{2i})^2 + \end{aligned} \quad (4.5)$$

$$+ (r_{31}x_{1i} + r_{32}x_{2i} + r_{33}x_{3i} + t_3 - y_{3i})^2.$$

Предположим, что $X = \{x_1, \dots, x_n\}$ лежит в плоскости-Охз или в плоскости-Оху. Также 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 решим вариационную задачу относительно матрицы R:

$$\frac{\partial J(R)}{\partial r_{1k}} = \sum_{i=1}^n 2 (r_{11}x_{1i} + r_{12}x_{2i} + r_{13}x_{3i} - y_{1i})x_{ki} = 0, k = 1, 2, 3. \quad (4.6)$$

Тогда

$$\sum_{i=1}^n (r_{1m}x_{mi} + r_{1n}x_{ni} - y_{1i})x_{ki} + r_{1k} \sum_{i=1}^n x_{ki}^2 = 0, \quad (4.7)$$

$$k, m, n = 1, 2, 3, m, n \neq k.$$

Из выражения (4.7) получим значение параметров матрицы поворота

$$r_{1k} = - \frac{\sum_{i=1}^n (r_{1m}x_{mi} + r_{1n}x_{ni} - y_{1i})x_{ki}}{\sum_{i=1}^n x_{ki}^2}. \quad (4.8)$$

Учитывая выражение (4.8), перепишем решение вариационной задачи в следующем виде:

$$\begin{aligned} J(R) = \sum_{i=1}^n \left(r_{1m}x_{mi} - \frac{\sum_{j=1}^n (r_{1m}x_{mj} + r_{1n}x_{nj} - y_{1j})x_{kj}}{\sum_{j=1}^n x_{kj}^2} x_{ki} + r_{1n}x_{ni} - y_{1i} \right)^2 + \\ + (r_{21}x_{1i} + r_{22}x_{2i} + r_{23}x_{3i} - y_{2i})^2 + \\ + (r_{31}x_{1i} + r_{32}x_{2i} + r_{33}x_{3i} - y_{3i})^2. \end{aligned} \quad (4.9)$$

Выполним преобразования в первом слагаемом функционала J(R) (4.9)

$$r_{1m}x_{mi} - \frac{\sum_{j=1}^n (r_{1m}x_{mj} + r_{1n}x_{nj} - y_{1j})x_{kj}}{\sum_{i=1}^n x_{ki}^2} x_{ki} + r_{1n}x_{ni} - y_{1i} = \quad (4.10)$$

$$= r_{1m} \left(x_{mi} - x_{ki} \frac{\sum_{j=1}^n x_{mj} x_{kj}}{\sum_{j=1}^n x_{kj}^2} \right) + r_{1n} \left(x_{ni} - x_{ki} \frac{\sum_{j=1}^n x_{nj} x_{kj}}{\sum_{j=1}^n x_{kj}^2} \right) -$$

$$- \left(y_{1i} - x_{ki} \frac{\sum_{j=1}^n y_{1j} x_{kj}}{\sum_{j=1}^n x_{kj}^2} \right).$$

Если подставить полученное выражение (4.10) в функционал $J(R)$, то получим

$$J(R) =$$

$$= \sum_{i=1}^n \left(r_{1m} \left(x_{mi} - x_{ki} \frac{\sum_{j=1}^n x_{mj} x_{kj}}{\sum_{j=1}^n x_{kj}^2} \right) + \right. \quad (4.11)$$

$$\left. + r_{1n} \left(x_{ni} - x_{ki} \frac{\sum_{j=1}^n x_{nj} x_{kj}}{\sum_{j=1}^n x_{kj}^2} \right) - \left(y_{1i} - x_{ki} \frac{\sum_{j=1}^n y_{1j} x_{kj}}{\sum_{j=1}^n x_{kj}^2} \right) \right)^2 +$$

$$+ (r_{21}x_{1i} + r_{22}x_{2i} + r_{23}x_{3i} - y_{2i})^2 +$$

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

Введем следующие определения:

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

$$G_{pi} = x_{ni} - x_{ki} \frac{\sum_{j=1}^n x_{nj} x_{kj}}{\sum_{j=1}^n x_{kj}^2}, \quad (4.12)$$

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

Используя введенные определения в формулах (4.12), представим функционал $J(R)$ в формуле (4.11) в следующем виде:

$$J(R) = \sum_{i=1}^n (r_{1m}G_{mi} + r_{1n}G_{pi} - G_{ki})^2 + (r_{21}x_{1i} + r_{22}x_{2i} + r_{23}x_{3i} - y_{2i})^2 + (r_{31}x_{1i} + r_{32}x_{2i} + r_{33}x_{3i} - y_{3i})^2. \quad (4.13)$$

Определим частную производную $J(R)$

$$\frac{\partial J(R)}{\partial r_{1m}} = 2 \sum_{i=1}^n (r_{1m} G_{mi} + r_{1n} G_{pi} - G_{ki}) G_{mi} = 0. \quad (4.14)$$

Далее получим формулу для вычисления r_{1m} :

$$\begin{aligned} \sum_{i=1}^n r_{1m} G_{mi}^2 + r_{1n} G_{mi} G_{pi} - G_{mi} G_{ki} &= r_{1m} \sum_{i=1}^n G_{mi}^2 + \\ &+ r_{1n} \sum_{i=1}^n G_{mi} G_{pi} - \sum_{i=1}^n G_{mi} G_{ki} = 0, \end{aligned} \quad (4.15)$$

$$r_{1m} = - \frac{r_{1n} \sum_{i=1}^n G_{mi} G_{pi} - \sum_{i=1}^n G_{mi} G_{ki}}{\sum_{i=1}^n G_{mi}^2}. \quad (4.16)$$

Подставим полученное решение в функционал $J(R)$

$$\begin{aligned} J(R) = \sum_{i=1}^n \left(- \frac{r_{1n} \sum_{j=1}^n G_{mj} G_{pj} - \sum_{j=1}^n G_{mj} G_{kj}}{\sum_{j=1}^n G_{mj}^2} G_{mi} + r_{1n} G_{pi} - G_{ki} \right)^2 + \\ + (r_{21} x_{1i} + r_{22} x_{2i} + r_{23} x_{3i} - y_{2i})^2 + (r_{31} x_{1i} + r_{32} x_{2i} + r_{33} x_{3i} - y_{3i})^2. \end{aligned} \quad (4.17)$$

Давайте выполним преобразования с элементами суммы в первом слагаемом в функционале $J(R)$ (4.17)

$$\begin{aligned} J(R) = \sum_{i=1}^n \left(- \frac{r_{1n} G_{mi} \sum_{j=1}^n G_{mj} G_{pj} - G_{mi} \sum_{j=1}^n G_{mj} G_{kj}}{\sum_{j=1}^n G_{mj}^2} + r_{1n} G_{pi} - G_{ki} \right)^2 + \\ + (r_{21} x_{1i} + r_{22} x_{2i} + r_{23} x_{3i} - y_{2i})^2 + (r_{31} x_{1i} + r_{32} x_{2i} + r_{33} x_{3i} - y_{3i})^2. \end{aligned} \quad (4.18)$$

Далее перегруппируем слагаемые в (4.18)

$$\begin{aligned} J(R) = \sum_{i=1}^n \left(r_{1n} \left(G_{pi} - \frac{G_{mi} \sum_{j=1}^n G_{mj} G_{pj}}{\sum_{j=1}^n G_{mj}^2} \right) - \left(G_{ki} - \frac{G_{mi} \sum_{j=1}^n G_{mj} G_{kj}}{\sum_{j=1}^n G_{mj}^2} \right) \right)^2 + \\ + (r_{21} x_{1i} + r_{22} x_{2i} + r_{23} x_{3i} - y_{2i})^2 + (r_{31} x_{1i} + r_{32} x_{2i} + r_{33} x_{3i} - y_{3i})^2. \end{aligned} \quad (4.19)$$

Введем следующие обозначения:

$$q_1 = G_{pi} - \frac{G_{mi} \sum_{j=1}^n G_{mj} G_{pj}}{\sum_{j=1}^n G_{mj}^2}, q_2 = G_{ki} - \frac{G_{mi} \sum_{j=1}^n G_{mj} G_{kj}}{\sum_{j=1}^n G_{mj}^2}. \quad (4.20)$$

С учетом выражений (4.20) перепишем формулу (4.19)

$$J(R) = \sum_{i=1}^n (r_{1n} q_1 - q_2)^2 + \\ + (r_{21} x_{1i} + r_{22} x_{2i} + r_{23} x_{3i} - y_{2i})^2 + (r_{31} x_{1i} + r_{32} x_{2i} + r_{33} x_{3i} - y_{3i})^2. \quad (4.21)$$

Теперь мы можем определить частную производную $J(R)$ по r_{1n}

$$\frac{\partial J(R)}{\partial r_{1n}} = 2 \sum_{j=1}^n (r_{1n} q_1 - q_2) q_1 = 0. \quad (4.22)$$

Тогда

$$r_{1n} = \frac{\sum_{k=1}^n q_1 q_2}{\sum_{k=1}^n q_1^2}. \quad (4.23)$$

При $\sum_{k=1}^n q_1^2 \neq 0$ уравнение (4.23) существует, тогда определим первую строку матрицы поворота:

$$r_{1m} = - \frac{r_{1n} \sum_{i=1}^n G_{mi} G_{pi} - \sum_{i=1}^n G_{mi} G_{ki}}{\sum_{i=1}^n G_{mi}^2}, \\ r_{1k} = - \frac{\sum_{i=1}^n (r_{1n} x_{mi} + r_{1n} x_{ni} - y_{1i}) x_{ki}}{\sum_{i=1}^n x_{ki}^2}, \quad (4.24) \\ r_{1n} = \frac{\sum_{k=1}^n q_1 q_2}{\sum_{k=1}^n q_1^2}.$$

Вторая строка матрицы поворота:

$$\begin{aligned}
r_{2m} &= -\frac{r_{2n} \sum_{i=1}^n G_{mi} G_{pi} - \sum_{i=1}^n G_{mi} G_{ki}}{\sum_{i=1}^n G_{mi}^2}, \\
r_{2k} &= -\frac{\sum_{i=1}^n (r_{2m} x_{mi} + r_{2n} x_{ni} - y_{2i}) x_{ki}}{\sum_{i=1}^n x_{ki}^2}, \\
r_{2n} &= \frac{\sum_{k=1}^n q_1 q_2}{\sum_{k=1}^n q_1^2}.
\end{aligned} \tag{4.25}$$

Третья строка матрицы поворота:

$$\begin{aligned}
r_{3m} &= -\frac{r_{3n} \sum_{i=1}^n G_{mi} G_{pi} - \sum_{i=1}^n G_{mi} G_{ki}}{\sum_{i=1}^n G_{mi}^2}, \\
r_{3k} &= -\frac{\sum_{i=1}^n (r_{3m} x_{mi} + r_{3n} x_{ni} - y_{3i}) x_{ki}}{\sum_{i=1}^n x_{ki}^2}, \\
r_{3n} &= \frac{\sum_{k=1}^n q_1 q_2}{\sum_{k=1}^n q_1^2}.
\end{aligned} \tag{4.26}$$

Определим элементы вектора параллельного переноса T через элементы матрицы поворота:

$$t_k = \frac{1}{n} \sum_{i=1}^n (y_{ki} - (r_{k1} x_{1i} + r_{k2} x_{2i} + r_{k3} x_{3i})) = 0, k = 1, 2, 3. \tag{4.27}$$

4.2 Комбинированное решение задачи точка-точка для аффинных преобразований в трехмерном пространстве

Известно, что сходимость и точность классических алгоритмов ICP, предложенных в работах Бесла П., Чена Ю., Хорна Б., может быть существенно улучшена. Предложенные методы ICP характеризуют следующие недостатки: во-первых, классические методы ICP не учитывают при анализе и обработке информации локальную форму поверхности вокруг каждой точки в трехмерном облаке, а во-вторых, известные методы регистрации данных на основе ICP обладают большой вычислительной сложностью, при этом самая затратная операция – поиск ближайших точек; в-третьих, результат решения вариационной задачи зависит от правильности выбора начального приближения. Последний недостаток метода ICP может быть устранен, для этого в исследовании используются особые точки, которые, как известно, сопоставляют кадры данных без указания параметров инициализации. Хорн предложил решение условной вариационной задачи для метрики точка-точка в замкнутой форме для ортогональных преобразований, в данной работе получено решение в замкнутой форме для аффинных преобразований, которое позволяет:

- осуществлять регистрацию не жестких объектов на трехмерной сцене;
- получать решение вариационной задачи ICP в замкнутой форме для различных вырожденных случаев, связанных с расположением точек в 3D облаке точек на одной прямой или в одной плоскости.

Последнее свойство может быть полезно при обнаружении или слежении за структурными элементами замкнутого трехмерного пространства. Для зданий и помещений данными элементами могут являться: потолок, стены, пол.

Предложенный в главе 3 алгоритм сопоставления данных [10] используется для обработки характерных точек на изображении в RGB-D кадре [10]. Процесс обработки особых точек представим в виде следующей последовательности шагов [26]:

1. Определение гистограмм ориентированных градиентов в RGB-D кадре (данные о цвете).
2. Сопоставление визуально связанных характеристик на двух последовательных изображениях.
3. Оценка параметров двумерной модели изображения на основе случайных выборок (RANSAC) с целью исключения выбросов в данных.
4. Решение вариационной задачи регистрации данных для визуально связанных характеристик изображения.

Рассмотрим более подробно п. 4. Решение вариационной задачи относительно визуально связанных характеристик изображения представим в виде

$$J(RV) = \frac{1}{|A_f|} \sum_{i \in A_f}^n w_i \|M(RV F_x^i) - M(F_y^i)\|^2, \quad (4.28)$$

где RV – матрица аффинного преобразования для визуальных характеристик сцены; w_i – весовые характеристики данных; F_x^i и F_y^i – особые точки в исходном и целевом кадре соответственно:

$$F_x^i = (x_{1f}^i, x_{2f}^i, x_{3f}^i)^T, F_y^i = (y_{1f}^i, y_{2f}^i, y_{3f}^i)^T. \quad (4.29)$$

M – функция преобразования координат точек трехмерной сцены F_x^i и $F_y^i \in \mathbb{R}^3$ в систему координат камеры $C^i = (C_x^i, C_y^i, D^i) \in \mathbb{R}^3$, где C_x^i, C_y^i – соответствующие координаты точек в пиксельном пространстве, D^i – значение глубины в пиксельном пространстве.

$$C_x^i = \frac{f}{x_{3f}^i} x_{1f}^i + O_x, C_y^i = \frac{f}{x_{3f}^i} x_{2f}^i + O_y, D^i = \sqrt{x_{1f}^{i2} + x_{2f}^{i2} + x_{3f}^{i2}}, \quad (4.30)$$

где O_x и O_y – координаты центра изображения в пиксельном пространстве, f – фокус камеры. Координаты визуально связанных характеристик в целевом кадре F_y^i могут быть получены в аналогичной манере.

Далее в параграфе 5.1 будет рассмотрено решение задачи извлечения семантической информации из 2D данных (изображения) и 3D данных (трехмерные облака точек). В исследовании используется детектор объектов (YOLO) и метод сегментации трехмерных данных (3-D Entangled Forests, 3DEF) для определения принадлежности объектов к классам предметной области (семантической маркировки окружающего пространства). Объединение результатов семантической сегментации методов 3DEF и YOLO осуществляется с использованием слияния на основе Байсевского подхода. После того, как будет получена реконструкция трехмерной сцены с семантической маркировкой пространства, становится возможным использовать полученные результаты о принадлежности объектов к классам предметной области для назначения весов соответствующим объектам классов. Назначение весов осуществляется на основе таблицы, в которой каждому классу поставлено в соответствие определенное значение веса. При этом каждой отдельной сцене может соответствовать своя таблица весов. Значения весов в таблицах устанавливаются экспертом. Такой подход позволяет для одних объектов сцены установить высокую значимость, а для других низкую. При этом важно, что фильтрация данных в трехмерных облаках точек осуществляется на основе информации о семантике объектов на трехмерной сцене.

Представим метод ICP с метрикой точка-точка для аффинных преобразований [31, 42] в виде следующей итерационной процедуры:

1. Предварительная обработка данных в одном или обоих облаках точек X и Y с целью уменьшения количества точек в анализируемых данных.
2. Установление соответствия между точками в X и Y . Для определения соответствующих точек используется k -d дерево.
3. Процедура взвешивания определенных пар соответствующих точек со стратегией назначения постоянных весов.
4. Процедура отклонения определенных пар соответствующих точек для исключения выбросов в данных (RANSAC метод или аналог).

5. Назначение метрики качества классификации для соответствующих пар точек, в работе используются метрика точка-точка.

6. Ключевой этап алгоритма ICP: нахождение решения вариационной задачи с использованием выбранной на шаге 5 метрики.

Вариационную задачу точка-точка для особых точек сцены и данных глубины можно представить в виде:

$$\begin{aligned} & \arg \min_{RV, RD} \arg \min J(RV, RD), \\ J(RV, RD) = & \alpha \frac{1}{W} \frac{1}{|A_f|} \sum_{i \in A_f}^m w_i \|M(RV F_x^i) - M(F_y^i)\|^2 + \\ & + (1 - \alpha) \frac{1}{W} \frac{1}{|A_d|} \sum_{j \in A_d}^n w_j \|RD x_j + T - y_j\|^2, \end{aligned} \quad (4.31)$$

где RV, RD – матрицы аффинного преобразования для особых точек и для данных о глубине сцены соответственно; α, W – весовые коэффициенты, подбираемые эмпирическим путем, которые позволяют управлять значимостью первого и второго слагаемых в функционале и нормировать данные соответственно; w_i, w_j – весовые коэффициенты, которые отражают семантические свойства объектов на трехмерной сцене (принадлежность объекта к определенному классу предметной области); 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} – значение геометрического преобразования метода ICP на основе оценки, полученной с использованием кинематической модели движения датчика; $k_{\max}$ – критерий останова по максимальному числу итераций в методе ICP; ε – критерий останова по значению ошибки минимизации E ; RT^* – лучшее (с точки зрения минимизации расстояния) аффинное преобразование для i -го шага; δ –

значение порога для inline точек. T_{km} используется в качестве значения начального геометрического преобразования, в том случае, если в RGB-D кадрах на изображениях не достаточно информации об особых точках для получения соответствующего геометрического преобразования.

Блок-схема метода для решения комбинированной задачи регистрации данных для аффинных преобразований представлена на рисунке 4.1. Представим детальное словесное описание данного алгоритма:

Шаг1. Вычислить ГНГ с использованием дескриптора ДЛ ГНГ для изображений в RGB-D кадрах.

Шаг 2. Выполнить алгоритм быстрого сопоставления данных на основе ориентированных градиентов для пары последовательных изображений, содержащих наборы особых точек F_x^i и F_y^i в виде двумерных облаков точек. Получить решение вариационной задачи регистрации данных для особых точек с использованием метрики точка-точка в виде геометрического преобразования RT^* и множества связей особых точек A_f .

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

Шаг 4. Определить соответствующие точки A_d в исходном и целевом облаке точек с использованием метода ближайших соседей. Выполнить процедуру взвешивания для определенных пар соответствующих точек и получить значения весовых коэффициентов для каждой полученной пары A_d . Значение веса устанавливается на основе результатов семантической разметки изображений в RGB-D кадрах. Если определенный объект на сцене не был отнесен ни к одному из классов, то для него устанавливается значение веса по умолчанию как для фонового объекта.

Шаг. 5 Решить комбинированную вариационную задачу с метрикой точка-точка для группы аффинных преобразований относительно преобразования RT^* , множества связей особых точек A_f , множества связей между трехмерными облаками точек A_d .

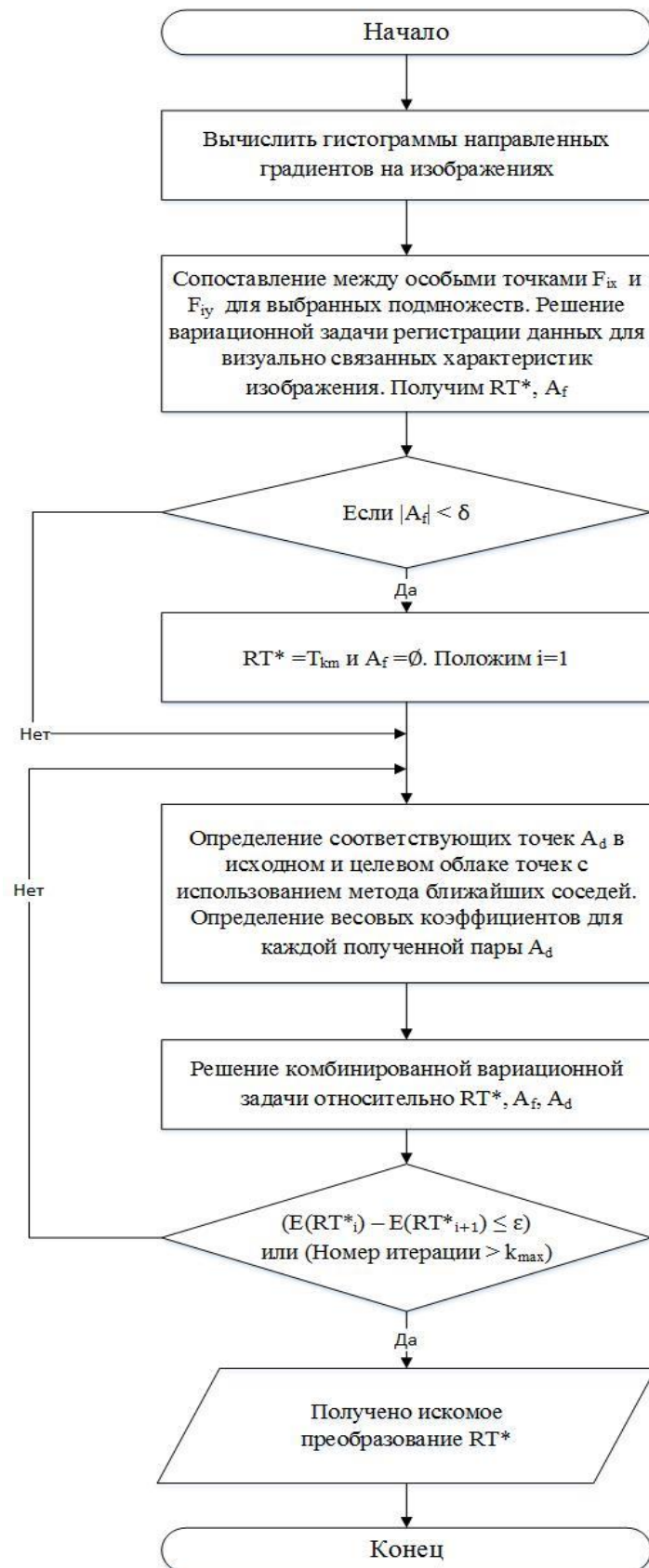


Рисунок 4.1 – Блок-схема алгоритма для решения комбинированной задачи регистрации данных для аффинных преобразований

Шаг 6. Выполнить проверку условия: $(E(RT^*_i) - E(RT^*_{i+1}) < \varepsilon)$ или (Номер итерации $> k_{\max}$). Если условие не верно, то необходимо вернуться на Шаг 4 данного метода и увеличить значение номера итераций $i=i+1$. Полученное на данном этапе геометрическое преобразование сохраняется с целью выполнения проверки условия (6) на следующем шаге. Если условие верно, то получено итоговое решение вариационной задачи – аффинное преобразование RT^* , итерации метода прекращаются.

Итеративная процедура завершается после того, как ошибка $E(RT^*_i)$ больше не уменьшается значительно или достигнуто максимальное число итераций. В противном случае вариационная задача решается относительно самого последнего преобразования.

4.3 Компьютерное моделирование

Подтвердим адекватность и достоверность предложенного комбинированного метода регистрации данных для аффинных преобразований результатами компьютерного моделирования для эталонных баз данных. Для проведения моделирования применялся компьютер на базе Intel Core i7 с графическим процессором. Оценим точность и сходимость предложенного метода на примере эталонной базы данных NYU Depth Dataset C2 [179]. Для этого выберем следующие тестовые наборы данных: Classrooms (классная комната), Living Rooms (1/4, жилая комната, фрагмент сцены, полученный с камеры с узким углом обзора), Offices (1/2, жилая комната, фрагмент сцены, полученный с камеры с широким углом обзора), Offices (2/2, офис №2, фрагмент сцены, полученный с камеры с широким углом обзора). Каждый тестовый набор данных содержит файлы дампа, изображение в формате ppm (Рис. 4.2, а), данные о глубине в формате pgm (Рис. 4.2, б), а также данные о семантической сегментации сцены. Для экспериментальных исследований использовалась камера Kinect 2.0 в качестве RGB-D сенсора. Данные для компьютерного моделирования (RGB-D кадры) были получены из разных положений мультисенсорных датчиков на сцене. Например, набор данных Classrooms содержит 688 ключевых кадров размером 640×480.

В работе проведено исследование точности и скорости сходимости предложенного комбинированного метода регистрации трехмерных данных и известных методов регистрации трехмерных данных на основе метода ICP. Были рассмотрены способы решения вариационной задачи ICP, использующие различные варианты минимизации функционала: с использованием метрики точка-точка с экстраполяцией [74] и без экстраполяции [71], с использованием метрики точка-плоскость [73]. Для проведения сравнительного анализа методов регистрации ICP по точности использована среднеквадратичная ошибка (RMSE).

Компьютерное моделирование по исследованию точности и сходимости проведено в условиях равномерного освещения сцены (набор данных Offices (1/2),

Сцена 1 с контролируемыми условиями) и в условиях неравномерного освещения сцены (набор данных Offices (1/2), Сцена 2 с неконтролируемыми условиями). В результате последовательности тестов для набора данных Offices (1/2) определен вид зависимости скорости сходимости методов регистрации от выбранного варианта минимизации функционала и условий освещенности сцены при проведении тестов (Рис. 4.3). Было установлено, что в контролируемых условиях комбинированный метод регистрации данных имеет график сходимости похожий на график сходимости метода, который использует метрику точка-плоскость (Рис. 4.3, а) в неконтролируемых условиях (при неравномерном освещении) предложенный метод показывает лучшую сходимость, чем указанные выше методы регистрации данных (Рис. 4.3, б).

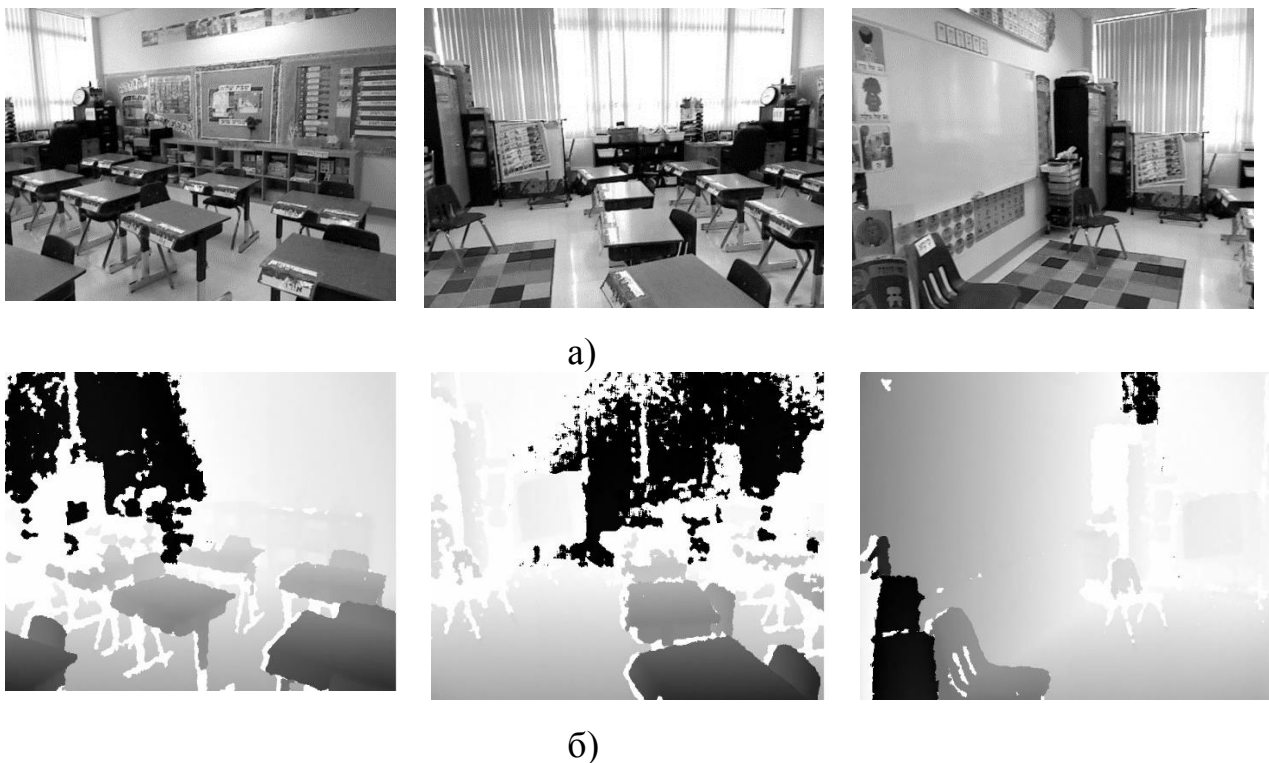


Рисунок 4.2 – Тестовый набор данных классная комната (Classrooms) из NYU Depth Dataset: а) визуальные данные кадра, б) данные глубины кадра

Дополнительно было установлено, что точность метода реконструкции зависит от количества особых точек в RGB-D кадре нелинейным образом – в виде функции с одним ярко выраженным пиком для всех типов дескрипторов.

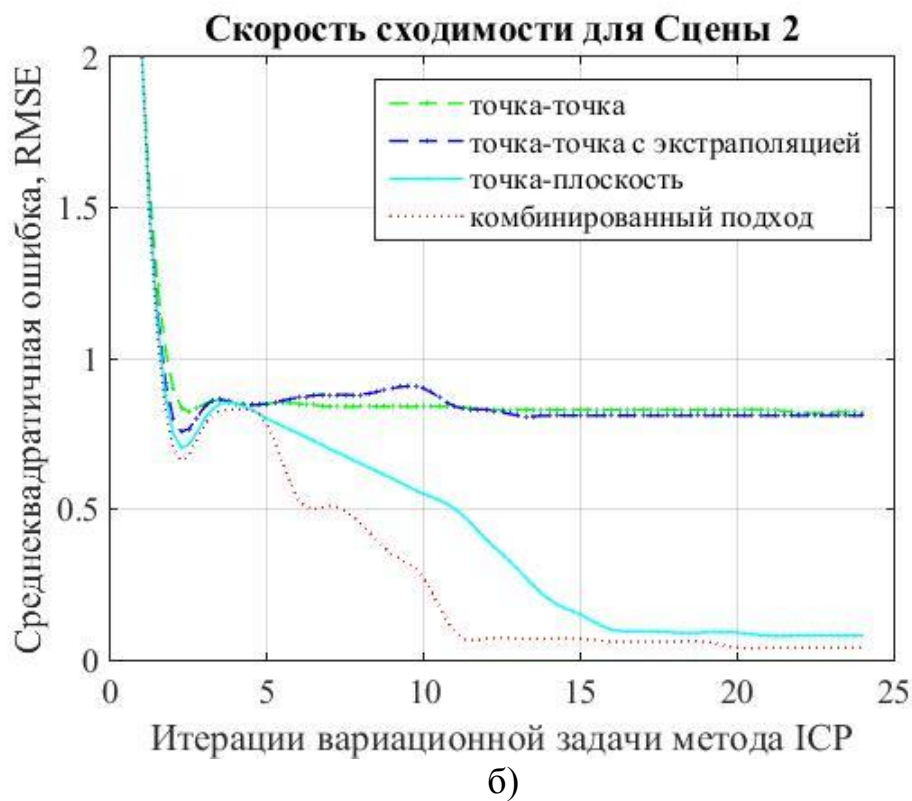
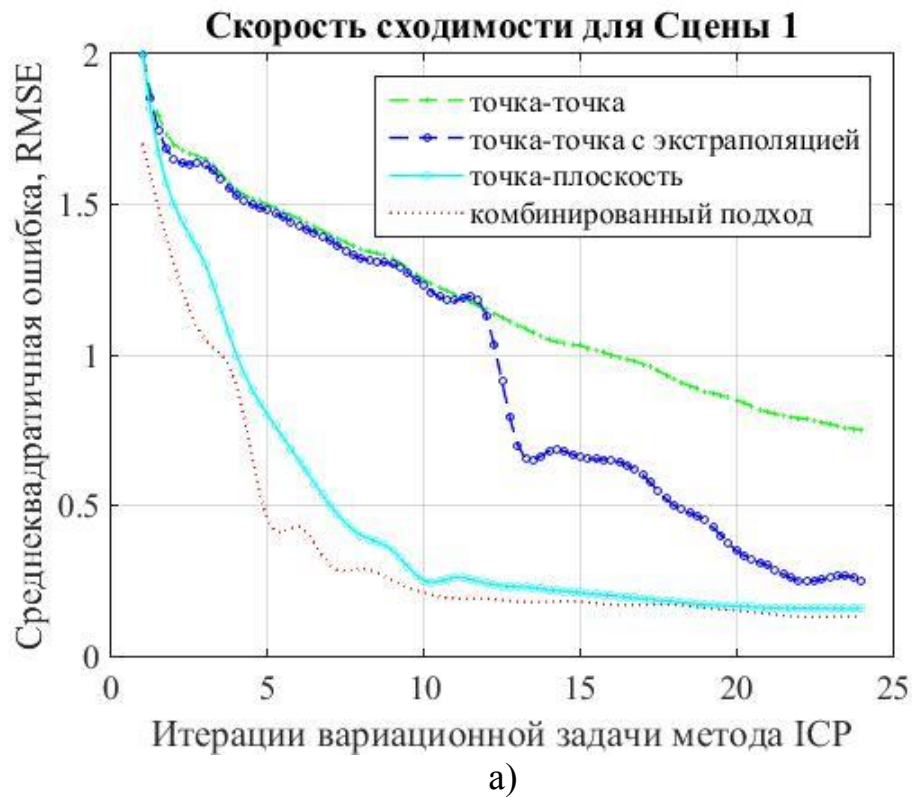


Рисунок 4.3 – Сравнение скорости сходимости в зависимости от ошибки метрики и условий наблюдения: а) в контролируемых условиях, б) в неконтролируемых условиях

Результаты реконструкции трехмерного объекта с использованием предложенного комбинированного метода регистрации данных с разных точек обзора на примере бюста головы человека представлены на рисунке 4.4.

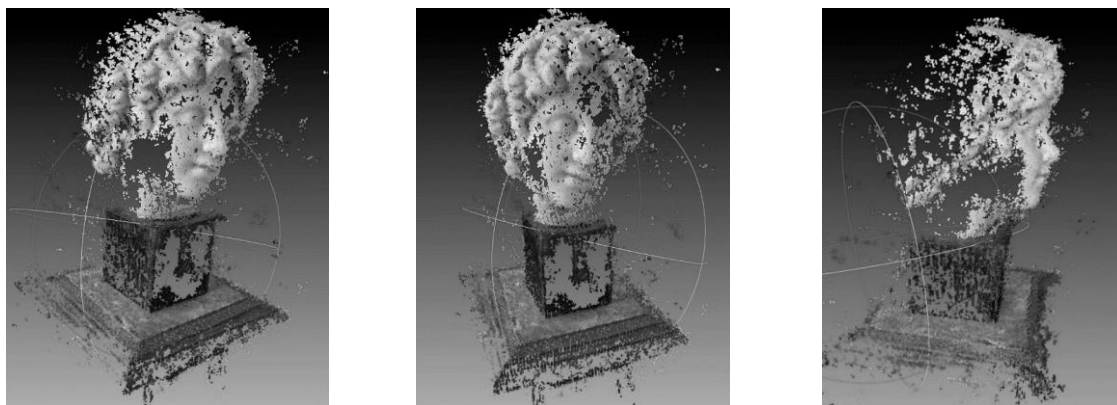
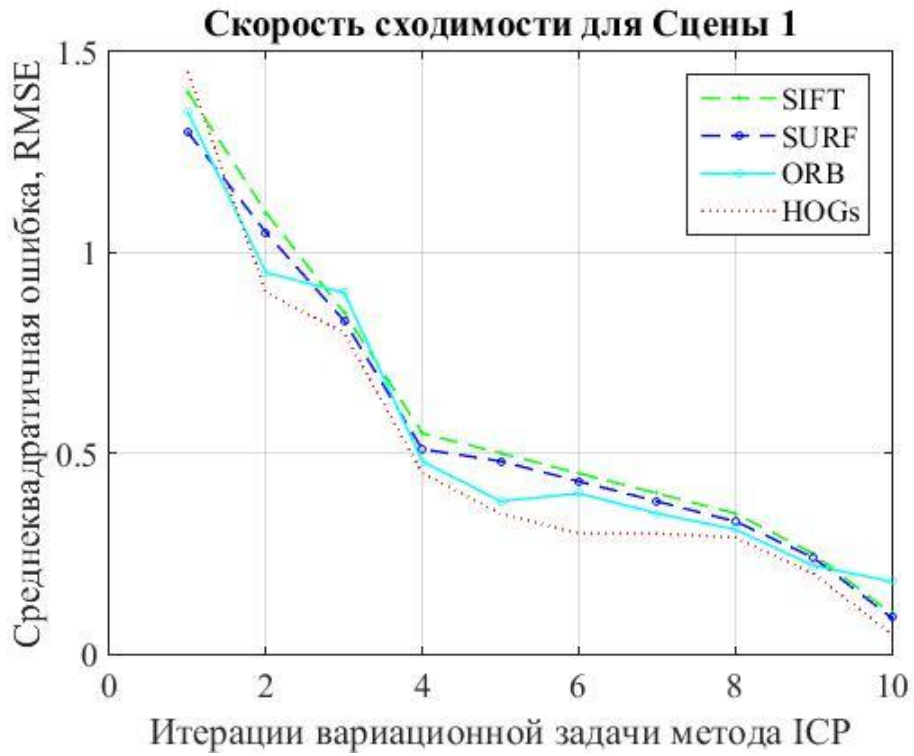


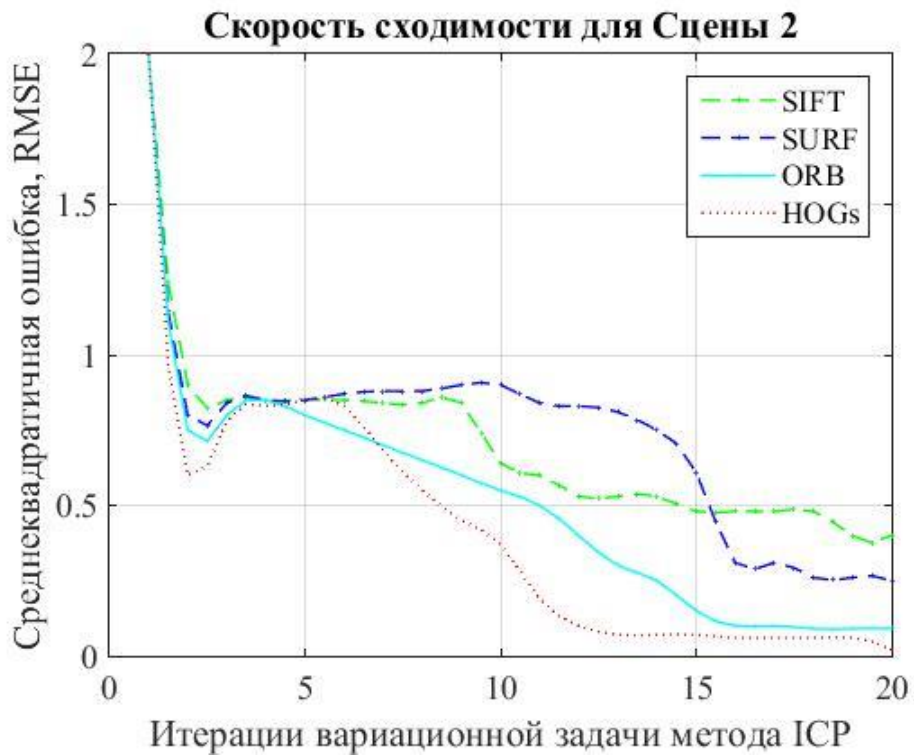
Рисунок 4.4 – Реконструкция трехмерного объекта на основе предложенного метода регистрации данных

В предложенном комбинированном методе регистрации для повышения точности решения вариационной задачи были применены данные об особых точках. Как известно, дескриптор HOGs (ДЛ ГНГ) инвариантен к неравномерному освещению и фотометрическим преобразованиям, которые проявляются на больших сценах [11, 174, 175]. Был проведен сравнительный анализ зависимости скорости сходимости методов регистрации от выбора типа дескриптора и условий проведения экспериментов (Рис. 4.5).

«Для ускорения процедуры сопоставления данных в работе используется пирамидальный подход, основанный на низкочастотной фильтрации и последующей дискретизации результирующего изображения/облака точек. Вычислительная сложность предлагаемого метода регистрации может быть оценена следующим образом: $n_1 + \frac{n_2 O_1}{F^1} + \dots + \frac{n_k O_1}{F^{k-1}}$, где k – количество шагов дискретизации, n_i – количество итераций в методе регистрации, выполненных на шаге k , O_1 – вычислительная сложность первого шага алгоритма, F – параметр, определяющий разбиение RGB-D кадра» [8, 13].



а)



б)

Рисунок 4.5 – Сравнение скорости сходимости в зависимости от типа дескриптора и условий наблюдения: а) в контролируемых условиях, б) в неконтролируемых условиях

4.4 Точное решение вариационной задачи ИСР для ортогональных преобразований

Важно учитывать вид истинного геометрического преобразования между парой трехмерных облаков точек: если преобразование изначально является ортогональным, то решение вариационной задачи ИСР для группы аффинных преобразований в большинстве случаев даст неверный результат. Верный результат может быть получен, если изначально соответствие между трехмерными облаками точек является близким к идеальному [175, 176]. Решение условной вариационной задачи минимизации функционала с ограничением на принадлежность решения классу ортогональных матриц приводит к необходимости работать с существенно нелинейным подмножеством ортогональных матриц многообразия матриц линейных преобразований. Поэтому рассматриваемая вариационная задача для ортогональных преобразований как математическая проблема значительно сложнее соответствующей задачи для аффинных преобразований. В работе разработан метод решения вариационной задачи в замкнутой форме для ортогональных преобразований, проведено сравнение разработанного метода с различными вариантами метода Хорна. Результаты компьютерного моделирования показали, что предложенный метод лучше стандартного метода регистрации, как по точности, так и по вычислительной сложности.

Пусть (Y) – поверхность, построенная из трехмерного облака точек Y , $(\cdot)$ – касательная плоскость к (Y) в заданной точке. Тогда вариационная задача точка-плоскость есть функция вида

$$J(R) = \sum_{i=1}^n (\langle Rx^i - y^i n^i \rangle)^2, \quad (4.32)$$

где $\langle \cdot, \cdot \rangle$ есть скалярное произведение, R – матрица аффинного преобразования в однородных координатах, $S(Y)$ – поверхность, построенная из облака точек Y , $T(y^i)$ – касательная плоскость к $S(Y)$ в точке y^i , n^i – унитарная нормаль к $T(y^i)$:

$$R = \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{pmatrix}, x^i = \begin{pmatrix} x_1^i \\ x_2^i \\ x_3^i \\ 1 \end{pmatrix}, y^i = \begin{pmatrix} y_1^i \\ y_2^i \\ y_3^i \\ 1 \end{pmatrix}, n^i = \begin{pmatrix} n_1^i \\ n_2^i \\ n_3^i \\ 1 \end{pmatrix}. \quad (4.33)$$

В работе [106] Лоу К.Л. привел графическую интерпретацию метрики точка-плоскость между двумя облаками точек (Рис. 4.6) в двумерном пространстве.

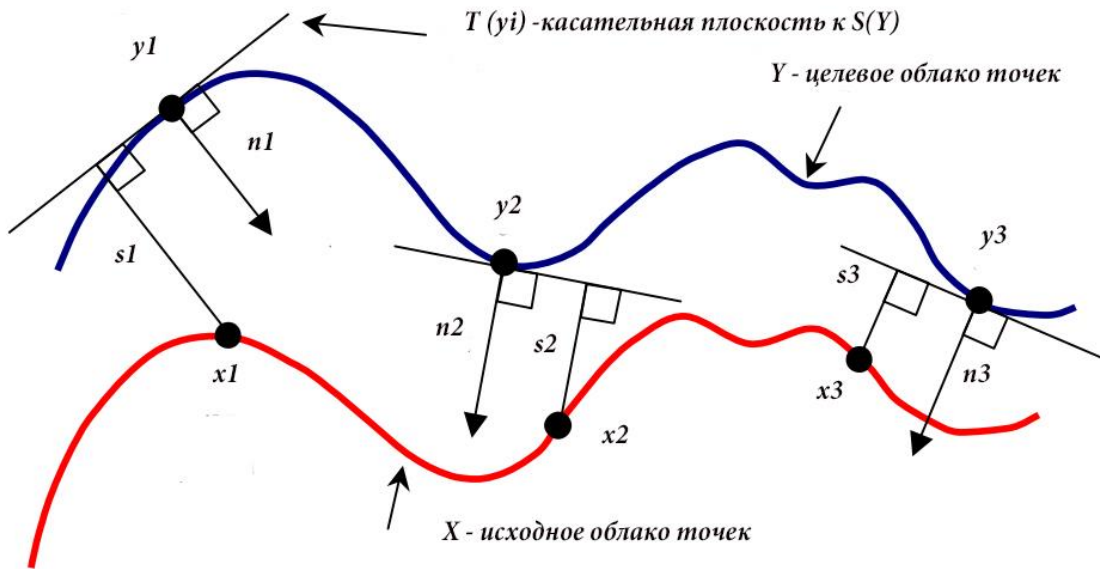


Рисунок 4.6 – Метрика точка-плоскость между двумя облаками точек

Тогда может быть получено решение поставленной вариационной задачи точка-плоскость, рассмотрим функцию $J(R)$ вида

$$\begin{aligned} J(R) &= \sum_{i=1}^n (\langle Rx^i - y^i, n^i \rangle)^2 = \sum_{i=1}^n (\langle Rx^i, n^i \rangle - \langle y^i, n^i \rangle)^2 = \\ &= \sum_{i=1}^n (\langle Rx^i, n^i \rangle)^2 - 2\langle Rx^i, n^i \rangle \langle y^i, n^i \rangle + (\langle y^i, n^i \rangle)^2. \end{aligned} \quad (4.34)$$

Слагаемое $\langle y^i, n^i \rangle$ является постоянным относительно R , следовательно, вариационная задача (4.34) принимает вид

$$\sum_{i=1}^n (\langle Rx^i, n^i \rangle)^2 - 2\langle Rx^i, n^i \rangle \langle y^i, n^i \rangle + (\langle y^i, n^i \rangle)^2 \rightarrow \min_R. \quad (4.35)$$

Рассмотрим компоненты градиента относительно переменной t :

$$\frac{\partial I}{\partial t_j} = \sum_{i=1}^n 2\langle Rx^i, n^i \rangle n_j^i - 2n_j^i \langle y^i, n^i \rangle = 0, j = 1, \dots, 3, \quad (4.36)$$

$$\sum_{i=1}^n \langle Rx^i, n^i \rangle n_j^i - n_j^i \langle y^i, n^i \rangle = \sum_{i=1}^n \langle Rx^i, n^i \rangle n_j^i - \sum_{i=1}^n n_j^i \langle y^i, n^i \rangle = 0. \quad (4.37)$$

Второе слагаемое в выражении (4.37) есть константна относительно x^i . Тогда выражение может быть выражено как

$$\sum_{i=1}^n \langle Rx^i, n^i \rangle n_j^i = c_j, j = 1, \dots, 3. \quad (4.38)$$

Далее обозначим через (XN) следующую матрицу

$$(XN)^i = \begin{pmatrix} x_1^i n_1^i & x_1^i n_2^i & x_1^i n_3^i & 0 \\ x_2^i n_1^i & x_2^i n_2^i & x_2^i n_3^i & 0 \\ x_3^i n_1^i & x_3^i n_2^i & x_3^i n_3^i & 0 \\ n_1^i & n_2^i & n_3^i & 0 \end{pmatrix}. \quad (4.39)$$

Затем рассмотрим операцию след матрицы для (XN)

$$\langle Rx^i, n^i \rangle = Tr(R \cdot (XN)^i), \quad (4.40)$$

где Tr – след матрицы.

На основе формулы (4.40) перепишем формулу (4.38)

$$\sum_{i=1}^n n_j^i Tr(R \cdot (XN)^i) = c_j, j = 1, \dots, 3. \quad (4.41)$$

Далее обозначим через $(n_j XN)$ множество матриц. Заметим, что выполняется следующие условие

$$n_j^i Tr(R \cdot (XN)^i) = Tr(R \cdot (n_j XN)^i). \quad (4.42)$$

С учетом условия в выражении (4.42) можно представить выражение (4.41) в виде выражения

$$\sum_{i=1}^n Tr(R \cdot (n_j XN)^i) = c_j. \quad (4.43)$$

Учтем свойство трассировки матрицы, тогда получим следующие зависимости

$$\sum_{i=1}^n Tr(R \cdot (n_j XN)^i) = Tr\left(\sum_{i=1}^n R \cdot (n_j XN)^i\right) = Tr\left(R \left(\sum_{i=1}^n (n_j XN)^i\right)\right) = c_j. \quad (4.44)$$

Аналогично найдем частные производные функционала $J(R)$ относительно r_{ij} , $i, j = 1, \dots, 3$:

$$\frac{\partial J}{\partial r_{ij}} = \sum_{k=1}^n 2\langle Rx^k, n^k \rangle x_j^k n_i^k - 2x_j^k n_i^k \langle y^k, n^k \rangle = 0, \quad (4.45)$$

$$i, j = 1, \dots, 3, k = 1, \dots, n.$$

Тогда

$$\begin{aligned} \sum_{k=1}^n \langle Rx^k, n^k \rangle x_j^k n_i^k - x_j^k n_i^k \langle y^k, n^k \rangle = \\ = \sum_{k=1}^n \langle Rx^k, n^k \rangle x_j^k n_i^k - \sum_{k=1}^n x_j^k n_i^k \langle y^k, n^k \rangle = 0. \end{aligned} \quad (4.46)$$

Следовательно, частная сумма $\sum_{k=1}^n x_j^k n_i^k \langle y^k, n^k \rangle$ есть константа, тогда

$$\sum_{k=1}^n x_j^k n_i^k \text{Tr}(R \cdot (XN)^k) = c_{ij}. \quad (4.47)$$

Обозначим через $x_j^k n_i^k XN^k$ набор матриц. Тогда выражение (4.47) примет вид

$$\sum_{k=1}^n \text{Tr}(R \cdot x_j^k n_i^k XN^k) = c_{ij}. \quad (4.48)$$

Таким образом, получен набор из 9 уравнений для $i, j = 1, \dots, 3$. Имеется множество из 2 линейных уравнений вида (4.44), (4.48) относительно 12 переменных r_{ij} , t_k , $i, j, k = 1, \dots, 3$, тогда решение может быть представлено в матричной форме как

$$M \cdot P = C, \quad (4.49)$$

где матрица M имеет размер 9×12 и состоит из элементов m_{ij} , C есть матрица из 12 столбцов c_i и c_{ij} , и P есть вектор вида $P = (r_{11}, r_{12}, r_{13}, t_1, r_{21}, r_{22}, r_{23}, t_2, r_{31}, r_{32}, r_{33}, t_3)$.

4.5 Комбинированное решение вариационной задачи ИСР для ортогональных преобразований

Итеративный алгоритм ближайших точек содержит два ключевых этапа:

- определение соответствующих точек в облаках точек;
- решение вариационной задачи: нахождение геометрического преобразования, которое совмещает два произвольно расположенных в пространстве облака точек наилучшим образом.

Представим общую схему алгоритма ИСР в виде итеративной процедуры следующего вида:

Шаг 1. Определение соответствующих точек в трехмерных данных $\forall x_k \in X, y_k = \{y \in Y \mid y(x_k, y) = \min\}$.

Шаг 2. Взвешивание определенных пар соответствующих точек со стратегией назначения постоянных весов w_k . Нахождение решения вариационной задачи (преобразование R и T) с использованием метрики точка-точка:

$$E(R, T) = \frac{1}{W} \sum_{N_p} w_k \|Rx_k + T - y_k\|^2, \quad W = \sum_{i=1}^{N_S} \sum_{j=1}^{N_T} w_{ij}, \quad (4.50)$$

где N_S и N_T – число точек в X и Y соответственно.

Шаг 3. Проверка условия

$$|E(R, T)_i - E(R, T)_{i+1}| < \varepsilon, \quad (4.51)$$

где i – номер итерации, ε – значение порога точности вычислений (RMSE). Преобразование (R, T) , вычисленное на Шаге 2 (4.50) будем применять к трехмерному облаку точек X до тех пор, пока условие (4.51) не верно. После окончания итераций будет найдено преобразование, при котором R и T минимизируют выражение (4.50), либо выход из итерационного процесса произойдет по одному из критериев останова работы алгоритма (см. раздел 4.3.)

В работе предложен комбинированный метод регистрации трехмерных данных для группы ортогональных преобразований

$$\underset{\lambda}{argmin} = \|J_F \lambda + R_F\|^2, \quad (4.52)$$

где $\lambda = [\varpi^T, v^T]^T$, $\varpi \in \mathbb{R}^3$ and, $v \in \mathbb{R}^3$, степень свободы 6D при движении мобильной платформы, J_F – функционал, который минимизирует расстояние для особых точек в RGB-D кадре, R_F – функционал, который минимизирует расстояние для данных о глубине в RGB-D кадре. В работе Вохминцева А. [8] показано, что для комбинированного решения вариационной задачи ICP функция ошибки может быть вычислена как

$$E = \alpha E_{ICP} + (1 - \alpha) E_{RGB-D}, \quad (4.53)$$

где α – весовой коэффициент, который определяет условный вес первого и второго слагаемого, а E_{ICP} , E_{RGB-D} – ошибки метода регистрации для трехмерных облаков точек и для изображений соответственно. Тогда, учитывая (4.31), функция совместной ошибки (4.53) для комбинированной вариационной задачи с метрикой точка-плоскость может быть вычислена как:

$$E(\hat{A}, T) = \underset{\hat{A}, T}{argmin} \left[\begin{aligned} & \alpha \frac{1}{w} \left(\frac{1}{|\mathbb{S}_f|} \sum_i w_i |\hat{A} P_X^i - P_Y^i|^2 \right) + \\ & + (1 - \alpha) \frac{1}{w} \left(\frac{1}{|\mathbb{S}_d|} \sum_j w_j \left| ((\hat{A} x_j + T - y_j) n_j) \right|^2 \right) \end{aligned} \right], \quad (4.54)$$

где функции P_X^i и P_Y^i в первом слагаемом вычисляют проекции особых точек (визуально связанных характеристик) из системы координат трехмерной сцены в систему координат RGB-D камеры, $\mathbb{S}^F$ – подмножество содержит связи между особыми точками в двух кадрах, $\mathbb{S}^D$ – подмножество содержит связи между соответствующими точками x_j и y_j в трехмерных облаках точек в двух кадрах данных, w_i и w_j – веса, которые связаны с семантическими характеристиками окружающей среды, n_j – унитарная нормаль, $\hat{A}$ – матрица геометрического преобразования поворот, T – вектор геометрического преобразования параллельный перенос. Тогда обозначим:

$$\hat{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, T = \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix}, x_i = \begin{pmatrix} x_j^1 \\ x_j^2 \\ x_j^3 \end{pmatrix}, y_i = \begin{pmatrix} y_j^1 \\ y_j^2 \\ y_j^3 \end{pmatrix}. \quad (4.55)$$

Представим матрицу геометрического преобразования, вектора облаков точек и вектор нормали в однородных координатах:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & t_1 \\ a_{21} & a_{22} & a_{23} & t_2 \\ a_{31} & a_{32} & a_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{pmatrix}, x_j = \begin{pmatrix} x_j^1 \\ x_j^2 \\ x_j^3 \\ 1 \end{pmatrix}, y_j = \begin{pmatrix} y_j^1 \\ y_j^2 \\ y_j^3 \\ 1 \end{pmatrix}, n_j = \begin{pmatrix} n_j^1 \\ n_j^2 \\ n_j^3 \\ 0 \end{pmatrix}. \quad (4.56)$$

Пусть $J(A)$ функционал вида

$$J(A) = \underset{A}{\operatorname{argmin}} \left[\alpha \frac{1}{W} \left(\frac{1}{|\mathbb{S}_f|} \sum_i w_i (\langle AP_{X'}^i, -P_{Y'}^i \rangle)^2 \right) + \right. \\ \left. + (1 - \alpha) \frac{1}{W} \left(\frac{1}{|\mathbb{S}_d|} \sum_j w_j (\langle Ax_j - y_j, n_j \rangle)^2 \right) \right]. \quad (4.57)$$

Далее рассмотрим отдельно второе слагаемое в функции (4.57)

$$\begin{aligned} J^R(A) &= \frac{1}{W} \left(\frac{1}{|\mathbb{S}_d|} \sum_j w_j (\langle Ax_j - y_j, n_j \rangle)^2 \right) = \\ &= \frac{1}{W} \left(\frac{1}{|\mathbb{S}_d|} \sum_j w_j (\langle Ax_j, n_j \rangle - \langle y_j, n_j \rangle)^2 \right) = \\ &= \frac{1}{W} \left(\frac{1}{|\mathbb{S}_d|} \sum_j w_j (\langle Ax_j, n_j \rangle^2 - 2 \langle Ax_j, n_j \rangle \langle y_j, n_j \rangle + \langle y_j, n_j \rangle^2) \right). \end{aligned} \quad (4.58)$$

Рассмотрим по отдельности компоненты градиента заданного функционала $\nabla J(A)$. Вначале решим комбинированную вариационную задачу ICP с метрикой точка-плоскость относительно вектора переноса T в замкнутой форме:

$$\frac{\partial J^R}{\partial t_m} = \frac{1}{W} \left(\frac{1}{|\mathbb{S}_d|} \sum_{j=1}^N 2w_j \langle Ax_j, n_j \rangle n_j^m - 2w_j n_j^m \langle y_j, n_j \rangle \right) = 0, m = 1, \dots, 3. \quad (4.59)$$

Аналогично получим решение комбинированной вариационной задачи ICP с метрикой точка-плоскость относительно матрицы поворота R

$$\frac{\partial J^R}{\partial a_{ml}} = \frac{1}{W} \left(\frac{1}{|\mathbb{S}_d|} \sum_{j=1}^N 2w_j \langle Ax_j, n_j \rangle x_j^l n_j^m - 2w_j x_j^l n_j^m \langle y_j, n_j \rangle \right) = 0, \quad (4.60)$$

$$m, l = 1, \dots, 3.$$

В работе [93] численными методами получено решение вариационной задачи ICP

$$E_{ICP} \approx \|\mathbf{J}_{ICP} \alpha + \mathbf{r}_{ICP}\|^2, \quad (4.61)$$

где $\mathbf{J}_{ICP}$, $\mathbf{r}_{ICP}$ – якобианы реконструкции трехмерных облаков точек.

Аналогично получим приближенное решение вариационной задачи E_{RGB-D} . Тогда, учитывая (4.58) и (4.61), для каждой итерации вариационной задачи (4.54) найдем решение уравнений:

$$(\mathbf{J}_{ICP}^T \mathbf{J}_{ICP} + \mathbf{J}_{RGB-D}^T \mathbf{J}_{RGB-D}) \lambda = \mathbf{J}_{ICP}^T \mathbf{r}_{ICP} + \mathbf{J}_{RGB-D}^T \mathbf{r}_{RGB-D}. \quad (4.62)$$

где $\mathbf{J}_{RGB-D}$, $\mathbf{r}_{RGB-D}$ – якобианы реконструкции для изображений.

4.6 Точное решение комбинированной вариационной задачи ICP для ортогональных преобразований

Введем следующие определения: $X = \{x_1, \dots, x_n\}$ – исходное трехмерное облако точек и $Y = \{y, \dots, y_m\}$ – целевое трехмерное облако точек в $\mathbb{R}^3$. В параграфе 2.2. было показано, что итеративный алгоритм ближайших точек (ICP) [71-73] может быть представлен как

$$Rx_i + T, \quad (4.63)$$

где R – матрица компонентов преобразования поворот, T – вектор компонентов преобразования перенос, $i = 1, \dots, n$.

Пусть $S(Y)$ – поверхность, построенная из облака точек Y , $T(y^i)$ – касательная плоскость к $S(Y)$ в точке y^i , n_j – есть унитарная (гильбертовая) нормаль к y^i , тогда представим вариационную задачу точка-плоскость алгоритма ICP как функцию:

$$E(R, T) = \operatorname{argmin}_{\mathbb{M}} \left[\alpha \frac{1}{W} \left(\frac{1}{|A_f|} \sum_i w_i |\widehat{M} F_x^i - F_y^i|^2 \right) + \right. \\ \left. + (1 - \alpha) \frac{1}{W} \left(\frac{1}{|A_d|} \sum_j w_j \left| ((\widehat{M} x_j - y_j) n_j) \right|^2 \right) \right]. \quad (4.64)$$

Представим вариационную задачу точка-плоскость в виде [37]

$$J(RV, RD) = \alpha \frac{1}{W} \frac{1}{|A_f|} \sum_{i=1}^n w_i (\langle RV * F_x^i - F_y^i n^i \rangle)^2 + \quad (4.65)$$

$$+ (1 - \alpha) \frac{1}{W} \frac{1}{|A_d|} \sum_{i=1}^n w_i (\langle RD * x^i - y^i n^i \rangle)^2,$$

где $\langle \cdot, \cdot \rangle$ есть скалярное произведение, RV – матрица аффинного преобразования в однородных координатах для визуальных характеристик сцены, RD – матрица аффинного преобразования в однородных координатах для данных о глубине сцены:

$$RD = \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{pmatrix}, RV = \begin{pmatrix} f_{11} & f_{12} & f_{13} & b_1 \\ f_{21} & f_{22} & f_{23} & b_2 \\ f_{31} & f_{32} & f_{33} & b_3 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (4.66)$$

Пусть x^i есть точка в облаке точек X , y^i есть унитарная (гильбертовая) нормаль к $T(y^i)$:

$$x^i = \begin{pmatrix} x_1^i \\ x_2^i \\ x_3^i \\ 1 \end{pmatrix}, y^i = \begin{pmatrix} y_1^i \\ y_2^i \\ y_3^i \\ 1 \end{pmatrix}, F_x^i = \begin{pmatrix} x_{1f}^i \\ x_{2f}^i \\ x_{3f}^i \\ 1 \end{pmatrix}, F_y^i = \begin{pmatrix} y_{1f}^i \\ y_{2f}^i \\ y_{3f}^i \\ 1 \end{pmatrix}, n^i = \begin{pmatrix} n_1^i \\ n_2^i \\ n_3^i \\ 1 \end{pmatrix}. \quad (4.67)$$

Обозначим $\eta_1 = \alpha \frac{1}{W} \frac{1}{|A_f|}$ и $\eta_2 = (1 - \alpha) \frac{1}{W} \frac{1}{|A_d|}$.

Тогда перепишем функционал $J(RV, RD)$

$$\begin{aligned} J(RV, RD) &= \eta_1 \sum_{i=1}^n w_i (\langle RV F_x^i - F_y^i n^i \rangle)^2 + \eta_2 \sum_{i=1}^n w_i (\langle RD x^i - y^i, n^i \rangle)^2 = \\ &= \eta_1 \sum_{i=1}^n w_i (\langle RV F_x^i, n^i \rangle - \langle F_y^i, n^i \rangle)^2 \\ &+ \eta_2 \sum_{i=1}^n w_i (\langle RD x^i, n^i \rangle - \langle y^i, n^i \rangle)^2 = \\ &= \eta_1 \sum_{i=1}^n (\langle RV F_x^i, n^i \rangle)^2 - 2 \langle RV F_x^i, n^i \rangle \langle F_y^i, n^i \rangle + (\langle F_y^i, n^i \rangle)^2 + \\ &+ \eta_2 \sum_{i=1}^n (\langle RD x^i, n^i \rangle)^2 - 2 \langle RD x^i, n^i \rangle \langle y^i, n^i \rangle + (\langle y^i, n^i \rangle)^2, \end{aligned} \quad (4.68)$$

где x_i точка в исходном облаке точек Y , n^i есть нормаль к $T(y^i)$.

Выражения $\langle F_y^i, n^i \rangle$ и $\langle y^i, n^i \rangle$ являются постоянными относительно RV и RD соответственно, тогда представим вариационную задачу (4.65) в виде

$$\begin{aligned}
& \eta_1 \sum_{i=1}^n (\langle RVF_x^i, n^i \rangle)^2 - 2\langle RVF_x^i, n^i \rangle \langle F_y^i, n^i \rangle + \\
& + \eta_2 \sum_{i=1}^n (\langle RDx^i, n^i \rangle)^2 - 2\langle RDx^i, n^i \rangle \langle y^i, n^i \rangle \rightarrow \min_{RV, RD}.
\end{aligned} \tag{4.69}$$

Выполним преобразования над скалярными произведениями $\langle RVF_x^i, n^i \rangle$ и $\langle RDx^i, n^i \rangle$ и представим их в следующем виде:

$$\begin{aligned}
\langle RVx^i, n^i \rangle = & (f_{11}F_{1x}^i n_1^i + f_{12}F_{2x}^i n_1^i + f_{13}F_{3x}^i n_1^i + b_1 n_1^i) + (f_{21}F_{1x}^i n_2^i + \\
& + f_{22}F_{2x}^i n_2^i + f_{23}F_{3x}^i n_2^i + b_2 n_2^i) + (f_{31}F_{1x}^i n_3^i + f_{32}F_{2x}^i n_3^i + f_{33}F_{3x}^i n_3^i + \\
& + b_3 n_3^i) + (0),
\end{aligned} \tag{4.70}$$

$$\begin{aligned}
\langle RDx^i, n^i \rangle = & (r_{11}x_1^i n_1^i + r_{12}x_2^i n_1^i + r_{13}x_3^i n_1^i + t_1 n_1^i) + (r_{21}x_1^i n_2^i + \\
& + r_{22}x_2^i n_2^i + r_{23}x_3^i n_2^i + t_2 n_2^i) + (r_{31}x_1^i n_3^i + r_{32}x_2^i n_3^i + r_{33}x_3^i n_3^i + t_3 n_3^i) + \\
& + (0).
\end{aligned} \tag{4.71}$$

Как было показано в параграфах 4.2 и 4.5 функционал $J(R)$ является функцией от 12 переменных. Тогда рассмотрим компоненты градиента $\nabla (RV, RD)$. Вначале рассмотрим компоненты градиента относительно переменных t и b :

$$\frac{\partial J}{\partial t_j} = \sum_{i=1}^n 2\langle RDx^i, n^i \rangle n_j^i - 2n_j^i \langle y^i, n^i \rangle = 0, j = 1, 2, 3 \tag{4.72}$$

и

$$\frac{\partial J}{\partial b_j} = \sum_{i=1}^n 2\langle RVF_x^i, n^i \rangle n_j^i - 2n_j^i \langle F_y^i, n^i \rangle = 0, j = 1, 2, 3. \tag{4.73}$$

Далее воспользуемся свойствами суммы:

$$\sum_{i=1}^n \langle RDx^i, n^i \rangle n_j^i - n_j^i \langle y^i, n^i \rangle = \sum_{i=1}^n \langle RDx^i, n^i \rangle n_j^i - \sum_{i=1}^n n_j^i \langle y^i, n^i \rangle = 0. \tag{4.74}$$

Сумма $\sum_{i=1}^n n_j^i \langle y^i, n^i \rangle$ является постоянной относительно RD, тогда обозначим сумму α_j следующим образом $\sum_{i=1}^n n_j^i \langle y^i, n^i \rangle = \alpha_j$.

Тогда с учетом данного замечания представим формулу (4.74) в следующем виде

$$\alpha_j = \sum_{i=1}^n \langle RD x^i, n^i \rangle n_j^i. \quad (4.75)$$

Аналогично сумма $\sum_{i=1}^n n_j^i \langle F_y^i, n^i \rangle$ является постоянной относительно RV, тогда

$$\beta_j = \sum_{i=1}^n \langle RV F_x^i, n^i \rangle n_j^i. \quad (4.76)$$

Далее обозначим через (XN) и (FXN) следующие матрицы соответственно:

$$(XN)^i = x^i n^i = \begin{pmatrix} x_1^i n_1^i & x_1^i n_2^i & x_1^i n_3^i & 0 \\ x_2^i n_1^i & x_2^i n_2^i & x_2^i n_3^i & 0 \\ x_3^i n_1^i & x_3^i n_2^i & x_3^i n_3^i & 0 \\ n_1^i & n_2^i & n_3^i & 0 \end{pmatrix} \quad (4.77)$$

и

$$(FXN)^i = F_x^i n^i = \begin{pmatrix} F_{1x}^i n_1^i & F_{1x}^i n_2^i & F_{1x}^i n_3^i & 0 \\ F_{2x}^i n_1^i & F_{2x}^i n_2^i & F_{2x}^i n_3^i & 0 \\ F_{3x}^i n_1^i & F_{3x}^i n_2^i & F_{3x}^i n_3^i & 0 \\ n_1^i & n_2^i & n_3^i & 0 \end{pmatrix}. \quad (4.78)$$

Заметим, что:

$$\langle RD x^i, n^i \rangle = Tr(RD \cdot (XN)^i) \quad (4.79)$$

и

$$\langle RV F_x^i, n^i \rangle = Tr(RV \cdot (FXN)^i), \quad (4.80)$$

где Tr – след матрицы. В работах [36, 37] представлено доказательство данного утверждения.

Давайте представим выражения (4.75) и (4.76) с учетом выражений (4.79) и (4.80) соответственно как:

$$\sum_{i=1}^n n_j^i Tr(RD \cdot (XN)^i) = \alpha_j, \quad \sum_{i=1}^n n_j^i Tr(RV \cdot (FXN)^i) = \beta_j. \quad (4.81)$$

Обозначим через $(n_j XN)$ и $(n_j FXN)$ следующие матрицы:

$$(n_j XN)^i = n_j^i (XN)^i = n_j^i \begin{pmatrix} x_1^i n_1^i & x_1^i n_2^i & x_1^i n_3^i & 0 \\ x_2^i n_1^i & x_2^i n_2^i & x_2^i n_3^i & 0 \\ x_3^i n_1^i & x_3^i n_2^i & x_3^i n_3^i & 0 \\ n_1^i & n_2^i & n_3^i & 0 \end{pmatrix} \quad (4.82)$$

и

$$(n_j FXN)^i = n_j^i (FXN)^i = n_j^i \begin{pmatrix} F_{1x}^i n_1^i & F_{1x}^i n_2^i & F_{1x}^i n_3^i & 0 \\ F_{2x}^i n_1^i & F_{2x}^i n_2^i & F_{2x}^i n_3^i & 0 \\ F_{3x}^i n_1^i & F_{3x}^i n_2^i & F_{3x}^i n_3^i & 0 \\ n_1^i & n_2^i & n_3^i & 0 \end{pmatrix}. \quad (4.83)$$

Заметим, что выполняются следующие условия:

$$n_j^i Tr(RD \cdot (XN)^i) = Tr(RD \cdot (n_j XN)^i) \quad (4.84)$$

и

$$n_j^i Tr(RV \cdot (FXN)^i) = Tr(RV \cdot (n_j FXN)^i). \quad (4.85)$$

С учетом условия в выражениях (4.84) и (4.85) можно представить выражения (4.81) в следующем виде:

$$\sum_{i=1}^n Tr(RD \cdot (n_j XN)^i) = \alpha_j \text{ и } \sum_{i=1}^n Tr(RV \cdot (n_j FXN)^i) = \beta_j. \quad (4.86)$$

На основе свойства трассировки матрицы получим в формулы:

$$\begin{aligned} \sum_{i=1}^n Tr(RD \cdot (n_j XN)^i) &= Tr\left(\sum_{i=1}^n RD \cdot (n_j XN)^i\right) = \\ &= Tr\left(RD \left(\sum_{i=1}^n (n_j XN)^i\right)\right) = \alpha_j \end{aligned} \quad (4.87)$$

и

$$\begin{aligned} \sum_{i=1}^n Tr(RV \cdot (n_j FXN)^i) &= Tr\left(\sum_{i=1}^n RV \cdot (n_j FXN)^i\right) = \\ &= Tr\left(RV \left(\sum_{i=1}^n (n_j FXN)^i\right)\right) = \beta_j. \end{aligned} \quad (4.88)$$

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

$$\sum_{i=1}^n (n_j XN)^i = SD^j, \quad \sum_{i=1}^n (n_j FXN)^i = SV^j, \quad (4.89)$$

где

$$SD^j = \begin{pmatrix} s_{11}^{jd} & s_{12}^{jd} & s_{13}^{jd} & 0 \\ s_{21}^{jd} & s_{22}^{jd} & s_{23}^{jd} & 0 \\ s_{31}^{jd} & s_{32}^{jd} & s_{33}^{jd} & 0 \\ s_{41}^{jd} & s_{42}^{jd} & s_{43}^{jd} & 0 \end{pmatrix}, \quad SV^j = \begin{pmatrix} s_{11}^{jv} & s_{12}^{jv} & s_{13}^{jv} & 0 \\ s_{21}^{jv} & s_{22}^{jv} & s_{23}^{jv} & 0 \\ s_{31}^{jv} & s_{32}^{jv} & s_{33}^{jv} & 0 \\ s_{41}^{jv} & s_{42}^{jv} & s_{43}^{jv} & 0 \end{pmatrix}. \quad (4.90)$$

Используя определения (4.90), запишем выражения (4.87) и (4.88) как:

$$Tr(RD \cdot SD^j) = \alpha_j \text{ и } Tr(RV \cdot SV^j) = \beta_j. \quad (4.91)$$

В работе [37] приведены детализованные выражения в виде операций над матрицами.

Тогда найдем частные производные функционала $J(RD)$ относительно r_{ij} , f_{ij} , $i, j = 1, 2, 3$, $k = 1, \dots, n$:

$$\frac{\partial J}{\partial r_{ij}} = \sum_{k=1}^n 2 \langle RD x^k, n^k \rangle x_j^k n_i^k - 2 x_j^k n_i^k \langle y^k, n^k \rangle = 0, \quad (4.92)$$

и

$$\frac{\partial J}{\partial f_{ij}} = \sum_{k=1}^n 2 \langle RV F_x^k, n^k \rangle F_{jx}^k n_i^k - 2 F_{jx}^k n_i^k \langle F_y^k, n^k \rangle = 0. \quad (4.93)$$

Тогда:

$$\begin{aligned} \sum_{k=1}^n \langle RD x^k, n^k \rangle x_j^k n_i^k - x_j^k n_i^k \langle y^k, n^k \rangle = \\ = \sum_{k=1}^n \langle RD x^k, n^k \rangle x_j^k n_i^k - \sum_{k=1}^n x_j^k n_i^k \langle y^k, n^k \rangle = 0, \end{aligned} \quad (4.94)$$

$$\begin{aligned} \sum_{k=1}^n \langle RV F_x^k, n^k \rangle F_{jx}^k n_i^k - F_{jx}^k n_i^k \langle F_y^k, n^k \rangle = \\ = \sum_{k=1}^n \langle RV F_x^k, n^k \rangle F_{jx}^k n_i^k - \sum_{k=1}^n F_{jx}^k n_i^k \langle F_y^k, n^k \rangle = 0. \end{aligned} \quad (4.95)$$

Частная сумма $\sum_{k=1}^n x_j^k n_i^k \langle y^k, n^k \rangle$ постоянна относительно RD, а частная сумма $\sum_{i=1}^n F_{jx}^k n_i^k \langle F_y^k, n^k \rangle$ постоянна относительно RV, тогда:

$$\sum_{k=1}^n x_j^k n_i^k \langle y^k, n^k \rangle = \alpha_{ij} \text{ и } \sum_{k=1}^n F_{jx}^k n_i^k \langle F_y^k, n^k \rangle = \beta_{ij}. \quad (4.96)$$

Используя определения (4.96), представим выражения (4.94) и (4.95) как:

$$\sum_{k=1}^n \langle RD x^k, n^k \rangle x_j^k n_i^k = \alpha_{ij}, \quad (4.97)$$

$$\sum_{k=1}^n \langle RV F_x^k, n^k \rangle F_{jx}^k n_i^k = \beta_{ij}. \quad (4.98)$$

Принимая во внимание выражение (4.77), выражение (4.97) может быть представлено как

$$\sum_{k=1}^n x_j^k n_i^k \text{Tr}(RD \cdot (XN)^k) = \alpha_{ij}, \quad (4.99)$$

и выражение (4.98) с учетом (4.78) может быть представлено как

$$\sum_{k=1}^n F_{jx}^k n_i^k \text{Tr}(RV \cdot (FXN)^k) = \beta_{ij}. \quad (4.100)$$

Обозначим через $x_j^k n_i^k XN^k$ и $F_{jx}^k n_i^k FXN^k$ следующие матрицы:

$$x_j^k n_i^k XN^k = x_j^k n_i^k \begin{pmatrix} x_1^k n_1^k & x_1^k n_2^k & x_1^k n_3^k & 0 \\ x_2^k n_1^k & x_2^k n_2^k & x_2^k n_3^k & 0 \\ x_3^k n_1^k & x_3^k n_2^k & x_3^k n_3^k & 0 \\ n_1^k & n_2^k & n_3^k & 0 \end{pmatrix} \quad (4.101)$$

и

$$F_{jx}^k n_i^k FXN^k = F_j^k n_i^k \begin{pmatrix} F_1^k n_1^k & F_1^k n_2^k & F_1^k n_3^k & 0 \\ F_2^k n_1^k & F_2^k n_2^k & F_2^k n_3^k & 0 \\ F_3^k n_1^k & F_3^k n_2^k & F_3^k n_3^k & 0 \\ n_1^k & n_2^k & n_3^k & 0 \end{pmatrix}. \quad (4.102)$$

Тогда выражения (4.99) и (4.100) принимают соответственно вид:

$$\sum_{k=1}^n \text{Tr}(RD \cdot x_j^k n_i^k XN^k) = \alpha_{ij} \quad (4.103)$$

и

$$\sum_{k=1}^n \text{Tr}(RV \cdot F_{jx}^k n_i^k FXN^k) = \beta_{ij}. \quad (4.104)$$

Далее, используя матрицы SD^{ij} и SV^{ij} , $i, j = 1, 2, 3$, получим выражения вида:

$$\sum_{k=1}^n x_j^k n_i^k X N^k = S D^{ij} \text{ и } \sum_{k=1}^n F_j^k n_i^k F X N^k = S V^{ij}. \quad (4.105)$$

В работе [37] приведены детализованные выражения (4.105) в виде операций над матрицами.

Далее в параграфе 5.1 будет рассмотрено решение задачи извлечения семантической информации из 2D данных (изображения) и 3D данных (трехмерные облака точек). В исследовании используется детектор объектов (YOLO) и метод сегментации трехмерных данных (3-D Entangled Forests, 3DEF) для определения принадлежности объектов к классам предметной области (семантической маркировки окружающего пространства). Объединение результатов семантической сегментации методов 3DEF и YOLO осуществляется с использованием слияния на основе Байсевского подхода. После того, как будет получена реконструкция трехмерной сцены с семантической маркировкой пространства, становится возможным использовать полученные результаты о принадлежности объектов к классам предметной области для назначения весов соответствующим объектам классов. Назначение весов осуществляется на основе таблицы, в которой каждому классу поставлено в соответствие определенное значение веса. При этом каждой отдельной сцене может соответствовать своя таблица весов. Значения весов в таблицах устанавливаются экспертом. Такой подход позволяет для одних объектов сцены установить высокую значимость, а для других низкую. При этом важно, что фильтрация данных в трехмерных облаках точек осуществляется на основе информации о семантике объектов на трехмерной сцене.

Представим метод ICP с метрикой точка-плоскость для ортогональных преобразований в виде следующей итерационной процедуры [37]:

1. Предварительная обработка данных в одном или обоих плотных облаках точек X и Y с целью уменьшения количества точек в анализируемых данных.

2. Установление соответствия между точками в X и Y. Для определения соответствующих точек используется k-d дерево.
3. Процедура взвешивания определенных пар соответствующих точек со стратегией назначения постоянных весов.
4. Процедура отклонения определенных пар соответствующих точек для исключения выбросов в данных (RANSAC метод или аналог).
5. Назначение метрики качества классификации для соответствующих пар точек, используются метрика точка-плоскость.
6. Проецирование элемента многообразия матриц линейных преобразований на подмногообразие ортогональных матриц.
7. Ключевой этап алгоритма ICP: нахождение решения вариационной задачи с использованием выбранной на шаге 5 метрики.

Тогда может быть получено множество из 24 линейных уравнений вида относительно 24 переменных $r_{ij}, f_{ij}, t_k, b_k, i, j, k = 1, \dots, 3$. Таким образом, можно представить решение комбинированной вариационной задачи для ортогональных преобразований в матричной форме:

$$\begin{cases} Tr(RD \cdot SD^j) = \alpha_j \\ Tr(RV \cdot SV^j) = \beta_j \\ \sum_{k=1}^n Tr(RD \cdot SD^{ij}) = \alpha_{ij} \Rightarrow Z \cdot \Omega = \theta, \\ \sum_{k=1}^n Tr(RV \cdot SV^{ij}) = \beta_{ij} \end{cases} \quad (4.106)$$

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} , и Ω есть вектор вида

$$\Omega = (r_{11}, r_{12}, r_{13}, t_1, r_{21}, r_{22}, r_{23}, t_2, r_{31}, r_{32}, r_{33}, t_3, f_{11}, f_{12}, f_{13}, b_1, f_{21}, f_{22}, f_{23}, b_2, f_{31}, f_{32}, f_{33}, b_3). \quad (4.107)$$

Множества A_f и A_d , матрицы RV и RD, вектора x_j, y_j, F_x^i и F_y^i , а также параметры для комбинированного метода регистрации для ортогональных

преобразований имеют такое же значение как в комбинированном методе регистрации для аффинных преобразований (см. формулу 4.66). Аналогично задаче регистрации ICP для аффинных преобразований будем искать совместное решение вариационной задачи для ортогональных преобразований для частного случая, когда $RV = RD = RT$. Пусть T_{km} – значение геометрического преобразования метода ICP на основе оценки, полученной с использованием кинематической модели движения датчика; k_{max} – критерий останова по максимальному числу итераций в методе ICP; ε – критерий останова по значению ошибки минимизации E ; RT^* – лучшее (с точки зрения минимизации расстояния) аффинное преобразование для i -го шага; δ – значение порога для inline точек. T_{km} используется в качестве значения начального геометрического преобразования, в том случае, если в RGB-D кадрах на изображениях не достаточно информации об особых точках для получения соответствующего геометрического преобразования.

Блок-схема алгоритма для решения комбинированной задачи регистрации данных с метрикой точка-плоскость для ортогональных преобразований представлена на рисунке 4.7. Приведем детальное словесное описание данного алгоритма [42]:

Шаг 1. Вычислить гистограммы направленных градиентов с использованием дескриптора ДЛ ГНГ для изображений в RGB-D кадрах.

Шаг 2. Выполнить алгоритм сопоставления данных на основе ориентированных градиентов для пары последовательных изображений, содержащих наборы особых точек F_x^i и F_y^i в виде двумерных облаков точек. Получить решение вариационной задачи регистрации данных для особых точек с использованием метрики точка-плоскость в виде геометрического преобразования RT^* и множества связей особых точек A_f .

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

Шаг 4. Определить соответствующие точки A_d в исходном и целевом облаке точек с использованием метода ближайших соседей.

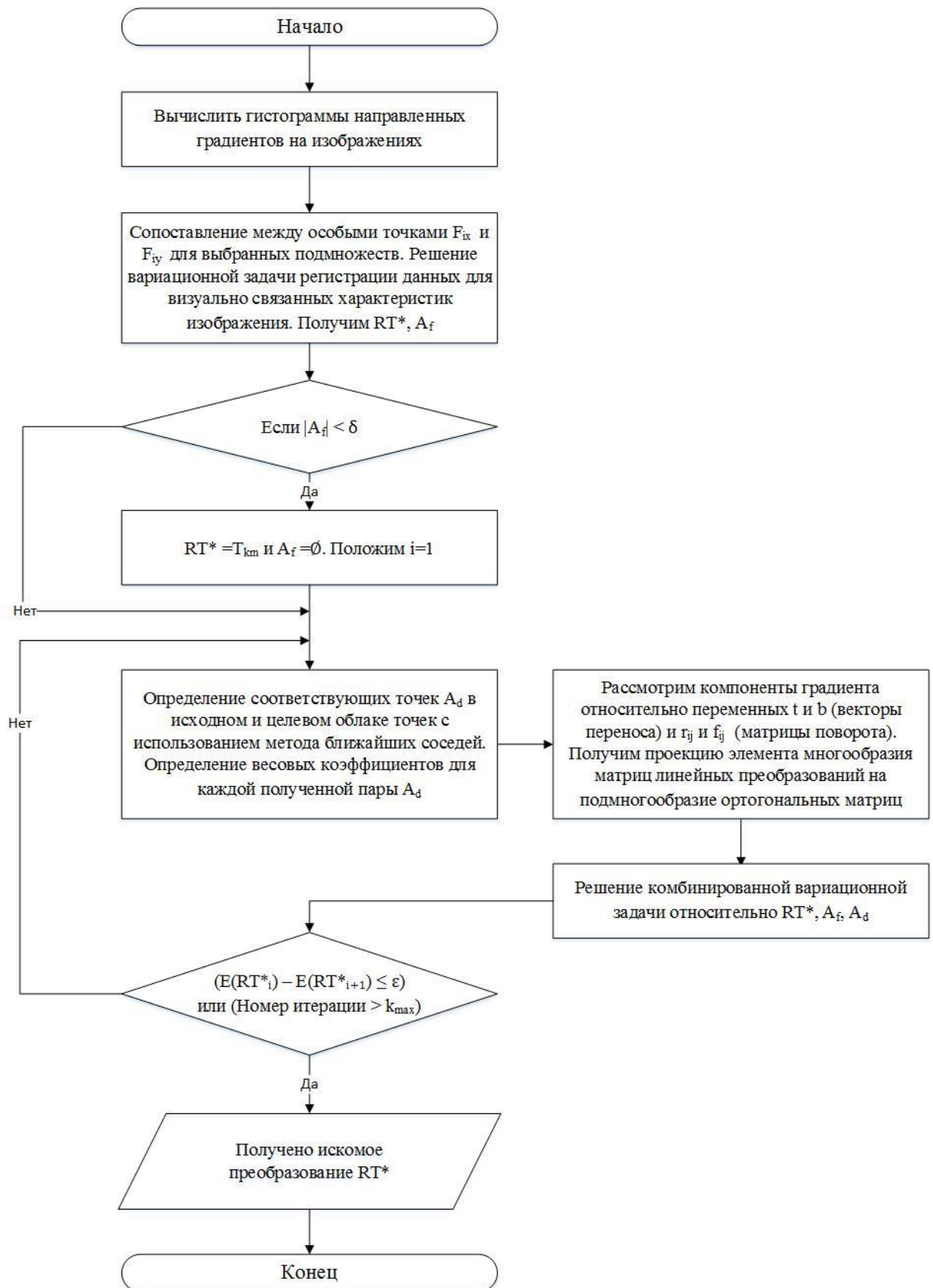


Рисунок 4.7 – Блок-схема алгоритма для решения комбинированной задачи регистрации данных для ортогональных преобразований

Выполнить процедуру взвешивания для определенных пар соответствующих точек и получить значения весовых коэффициентов для каждой полученной пары A_d . Значение веса устанавливается на основе результатов семантической разметки изображений в RGB-D кадрах. Если определенный объект на сцене не был отнесен ни к одному из классов, то для него устанавливается значение веса по умолчанию как для фонового объекта.

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 подмногообразие ортогональных матриц $F_{jx}^k n_i^k F X N^k$, $x_j^k n_i^k X N^k$ и $n_j^i X N$, $n_j^i F X N$.

Шаг 6. Решить комбинированную вариационную задачу с метрикой точка-плоскость для группы ортогональных преобразований относительно преобразования RT^* , множества связей особых точек A_f , множества связей между трехмерными облаками точек A_d .

Шаг 7. Выполнить проверку условия: $(E(RT^*_i) - E(RT^*_{i+1})) < \varepsilon$ или (Номер итерации $> k_{\max}$). Если условие не верно, то необходимо вернуться на Шаг 4 данного метода и увеличить значение номера итераций $i=i+1$. Полученное на данном этапе геометрическое преобразование сохраняется с целью выполнения проверки условия (7) на следующем шаге. Если условие верно, то получено итоговое решение вариационной задачи – преобразование RT^* , итерации метода прекращаются.

Итеративная процедура завершается после того, как ошибка $E(RT^*_i)$ больше не уменьшается значительно или достигнуто максимальное число итераций. В противном случае вариационная задача решается относительно самого последнего преобразования.

Шаг 5 обладает большой вычислительной сложностью, поэтому в общем случае решение комбинированной задачи для аффинных преобразований является менее затратной процедурой.

4.7 Компьютерное моделирование

Подтвердим адекватность и достоверность предложенного комбинированного метода регистрации данных для ортогональных преобразований результатами компьютерного моделирования для эталонных баз данных. Для проведения моделирования применялся компьютер на базе Intel Core i7 с графическим процессором. Оценим точность и сходимость предложенного метода на примере эталонной базы данных NYU Depth Dataset V2 [179]. Для этого выберем следующие тестовые наборы данных: Classrooms (классная комната), Living Rooms (1/4, жилая комната, фрагмент сцены, полученный с камеры с узким углом обзора), Offices (1/2, жилая комната, фрагмент сцены, полученный с камеры с широким углом обзора), Offices (2/2, офис №2, фрагмент сцены, полученный с камеры с широким углом обзора). Для экспериментальных исследований использовалась камера Kinect 2.0 в качестве RGB-D сенсора.

В работе проведено исследование точности и скорости сходимости предложенного комбинированного метода регистрации трехмерных данных и известных методов регистрации трехмерных данных на основе метода ICP. Были рассмотрены способы решения вариационной задачи ICP, использующие различные варианты минимизации функционала: точка-точка (метод Хорна) [71, 100], точка-точка с экстраполяцией (метод Хорна с экстраполяцией) [74, 101]. Для проведения сравнительного анализа методов регистрации ICP по точности использована среднеквадратичная ошибка (RMSE). Компьютерное моделирование по исследованию точности и сходимости проведено в условиях равномерного освещения сцены (набор данных Offices (1/2), Сцена 1 с контролируемыми условиями) и в условиях неравномерного освещения сцены (набор данных Offices (1/2), Сцена 2 с неконтролируемыми условиями). В результате компьютерного моделирования определена зависимость скорости сходимости методов трехмерной регистрации данных от выбора метрики и условий проведения экспериментов (Рис. 4.8).

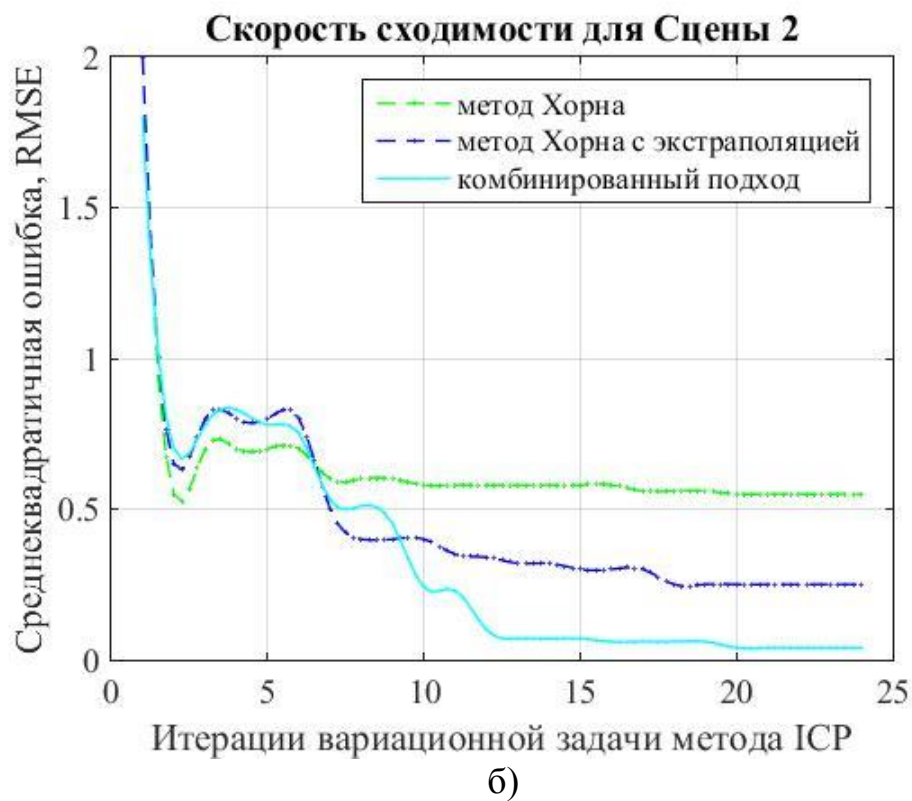
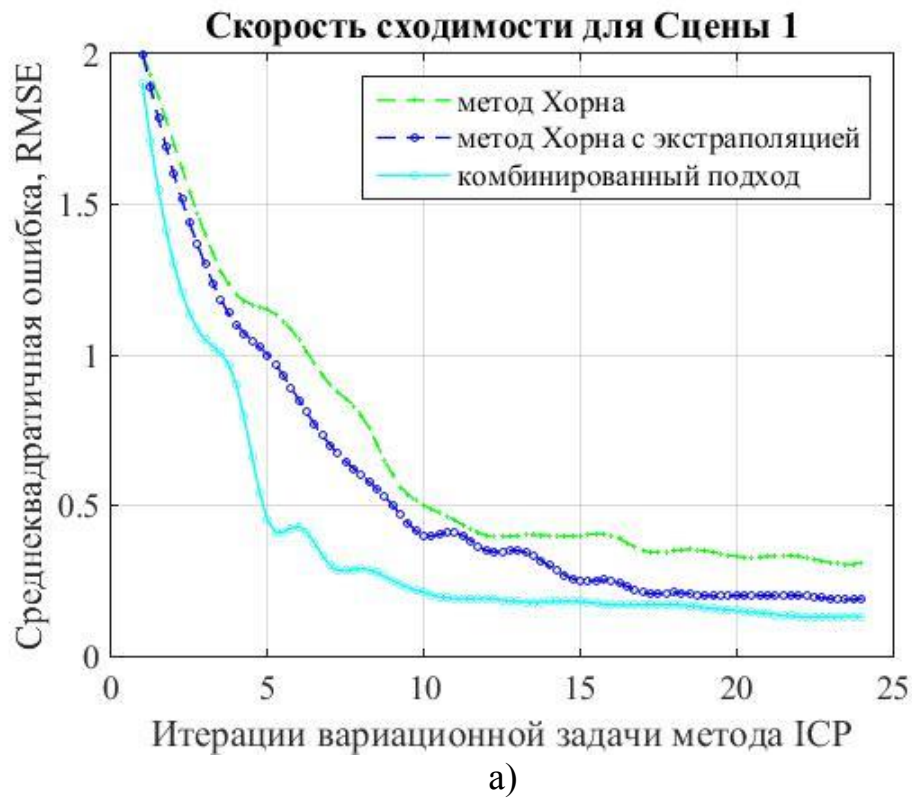


Рисунок 4.8 – Сравнение скорости сходимости в зависимости от ошибки метрики и условий наблюдения: а) в контролируемых условиях, б) в неконтролируемых условиях

При равномерном освещении предложенный метод регистрации трехмерных облаков точек имеет график сходимости похожий на график сходимости метода, который использует метрику точка-точка с экстраполяцией (Рис. 4.8, а) в неконтролируемых условиях (при неравномерном освещении) предложенный метод сходится на 1-2 итерацию позже, чем указанные выше методы регистрации данных (Рис. 4.8, б), но при этом показывает лучшую точность.

В работе проведено исследование точности и скорости сходимости методов регистрации трехмерных данных для ортогональных преобразований в зависимости от выбора типа дескриптора и условий проведения экспериментов (Рис. 4.9). Показано, что при равномерном освещении тип используемого дескриптора имеет ограниченное влияние на сходимость метода регистрации данных (Рис. 4.9, а), а при неравномерном освещении «дескриптор HOGs (ДЛ ГНГ) позволяет получить 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 (Рис. 4.9, б)» [37].

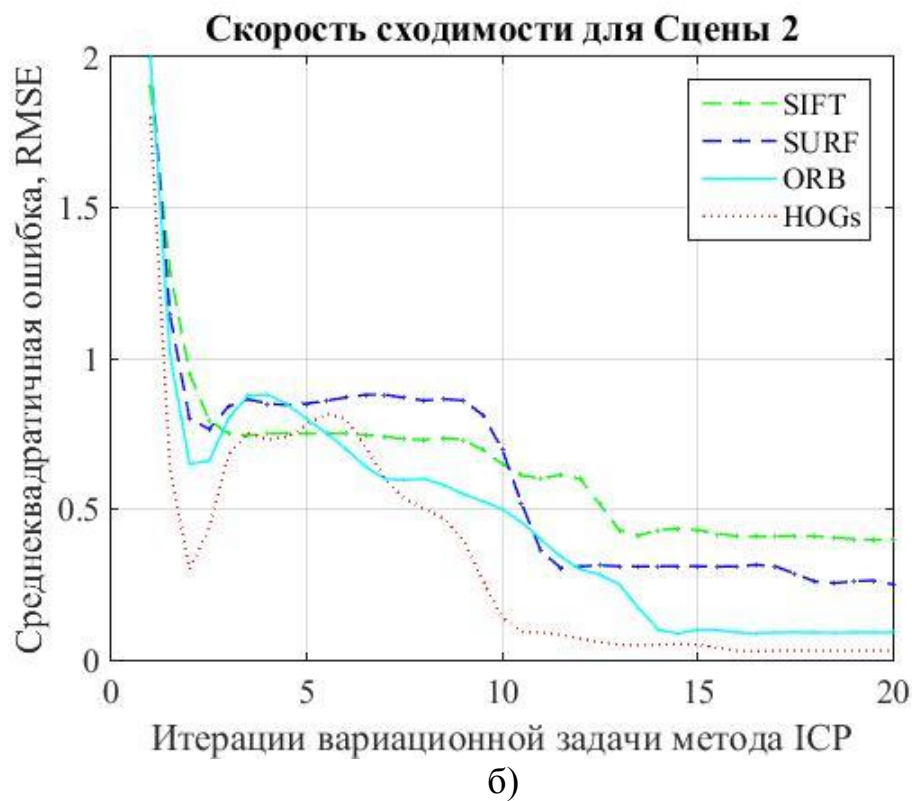
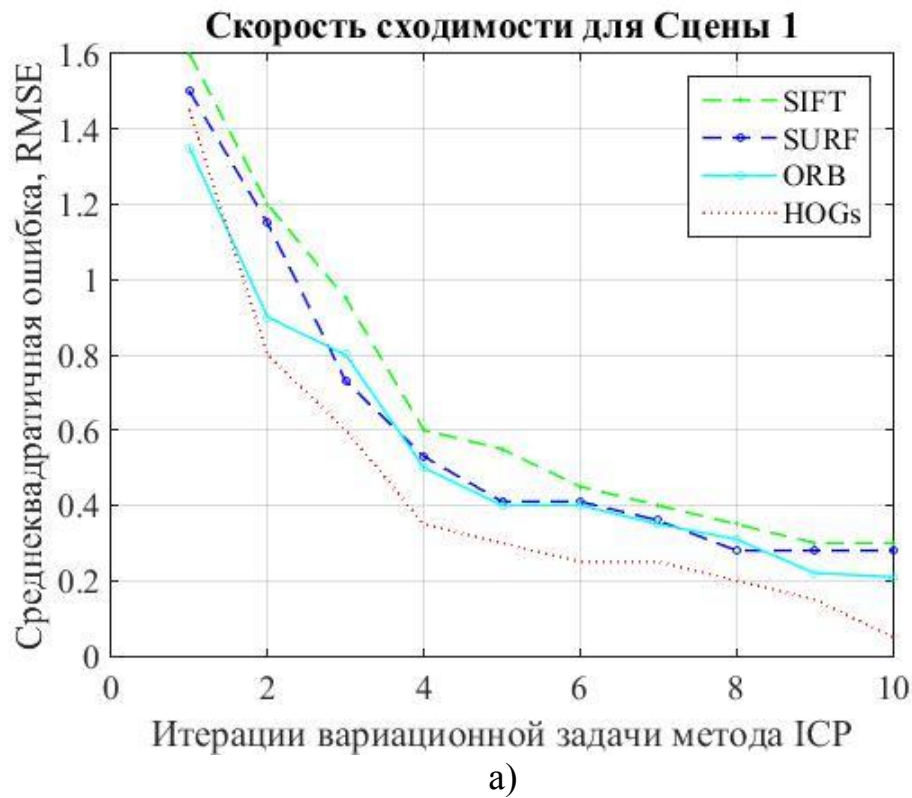


Рисунок 4.9 – Сравнение скорости сходимости в зависимости от выбора типа дескриптора и условий наблюдения: а) в контролируемых условиях, б) в неконтролируемых условиях

4.8 Выводы по главе

1. В работе проведен сравнительный анализ известных алгоритмов регистрации относительно облака трехмерных точек на основе итеративного алгоритма ближайших точек, было показано, что сходимость и точность известных итеративных методов и алгоритмов регистрации данных может быть существенно улучшена. Предложенные методы ICP характеризуют следующие недостатки: во-первых, классические методы ICP не учитывают при анализе и обработке информации локальную форму поверхности вокруг каждой точки в трехмерном облаке, а во-вторых, известные методы регистрации данных на основе ICP обладают большой вычислительной сложностью, при этом самая затратная операция – поиск ближайших точек.

2. Точность и сходимость метода ICP зависит от выбора метрики для решения вариационной задачи, наибольшее распространение при трехмерной реконструкции сцен получили метрика точка-точка (point-to-point) и метрика точка-плоскость (point-to-plane). Отличие методов ICP, использующих данные метрики для минимизации функционала, заключается в том, что в методе с метрикой точка-плоскость минимизируются расстояния от точек шаблона до касательных плоскостей к соответствующим точкам целевой поверхности, а в методе с метрикой точка-точка минимизируются расстояния между соответствующими точками. В современных приложениях робототехники, где требуется реконструировать крупномасштабные сцены в реальном масштабе времени, применение метрики точка-плоскость при решении вариационной задачи ICP позволяет добиться как лучшей сходимости, так и более высокой точности при сопоставлении трехмерных облаков точек. Однако методы с метрикой точка-плоскость имеют свои недостатки, т.к. дополнительно устанавливают требования к реконструируемой трехмерной сцене по числу геометрических отношений.

3. Итеративный алгоритм ближайших точек имеет два ключевых шага: определение соответствующих точек между парой произвольно расположенных в

пространстве трехмерных облаков точек и решение вариационной задачи с использованием выбранной метрики, которая находит лучшее геометрическое преобразование (с точки зрения минимизации расстояния между облаками точек). В исследовании предложены комбинированные методы ICP, которые улучшают качество работы двух ключевых шагов метода. Для решения задачи нахождения соответствия между точками двух облаков использован метод, основанный на геометрическом подобии для различных подмножеств в облаках точек. Некоторые результаты, касающиеся использования информации о характерных точках для решения вариационной задачи, были известны и ранее, однако явные формулы в поставленной постановке задачи для решения проблемы SLAM были впервые получены в рамках настоящей работы.

4. Вариационная задача алгоритма ICP, предложенного Хорном Б., находит оптимальное ортогональное преобразование, совмещающее два облака точек. В рамках работы вариационная задача алгоритма ICP распространена на класс аффинных преобразований. Проведенные вычислительные эксперименты показали значительное повышение точности комбинированного метода ICP для разного класса сцен.

5. Предложен метод решения вариационной задачи точка-плоскость для класса ортогональных преобразований без требования начальной инициализации. Было проведено сравнение предлагаемого метода регистрации данных с методом Хорна для метрики точка-точка с экстраполяцией и без экстраполяции. В результате компьютерного моделирования было установлено, что применение визуально связанных характеристик для решения вариационной задачи метода ICP позволяет улучшить точность и сходимость метода в контролируемых и неконтролируемых условиях.

6. Предложен метод решения вариационной задачи точка-точка в замкнутой форме для аффинных преобразований с использованием визуально связанных характеристик. Предложенный комбинированный метод решения вариационной задачи точка-точка в замкнутой форме для аффинных преобразований создает основу для применения метода Хорна на случай с не

ригидными объектами на сцене. В работе проведено исследование точности и скорости сходимости предложенного комбинированного метода регистрации трехмерных данных и известных методов регистрации трехмерных данных на основе метода ICP. Были рассмотрены способы решения вариационной задачи ICP, использующие различные варианты минимизации функционала: с использованием метрики точка-точка с экстраполяцией и без экстраполяции, с использованием метрики точка-плоскость. В результате компьютерного моделирования было установлено, что применение визуально связанных характеристик для решения вариационной задачи алгоритма ICP позволяет улучшить сходимость метода при неравномерном освещении.

7. Разработан комбинированный метод реконструкции трехмерной карты неизвестной окружающей среды с использованием доступной мультисенсорной информации для крупномасштабных сцен с контролируемыми и неконтролируемыми условиями.

8. Основные результаты по данной главе опубликованы в работах [8, 13, 26, 31, 33, 34, 36, 37, 42].

ГЛАВА 5 МЕТОД РЕШЕНИЯ ЗАДАЧИ НАВИГАЦИИ И ОПРЕДЕЛЕНИЯ ПОЛОЖЕНИЯ МОБИЛЬНОЙ ПЛАТФОРМЫ НА ОСНОВЕ КОМБИНИРОВАННОЙ МЕТОДОЛОГИИ SLAM

В данной главе рассмотрен метод определения динамического положения мобильной платформы в неизвестном пространстве на основе расширенной Калмановской фильтрации и комбинировании визуальной и семантической информации о трехмерной сцене. На каждом шаге предложенного метода выполняется две процедуры: вычисление сдвига положения мобильной платформы на трехмерной сцене и предсказание новой позиции платформы на сцене. Также в каждом RGB-D кадре можно выделить положение визуально связанных характеристик окружающей среды и определить позицию мобильной платформы на трехмерной сцене относительно них. Используем разницу между полученными оценками положений мобильной платформы для вычисления новых значений вероятности (веса) для всех визуально связанных характеристик в последовательности RGB-D кадров и для коррекции положения и траектории движения мобильной платформы на трехмерной сцене. Рассмотрим современное состояние проблемы. Для решения данной задачи в исследовании используется метод, основанный на EKF SLAM (Extended Kalman Filter) [53] и ориентации в пространстве на основе семантических маркеров. Входной информацией для каждой итерации метода EKF SLAM является множество неточных измерений положения мобильной платформы на сцене для некоторого дискретного момента времени k , выходной информацией для каждой итерации является последовательность оценок положения мобильной платформы на сцене. В EKF SLAM модель процесса получения точных и непрерывно обновляемых оценок вектора состояния из последовательности зашумленных измерений положения мобильной платформы принято представлять в виде системы алгебраических уравнений, которая сводится к матричному уравнению.

На сегодняшний день разработано множество точных успешных методов решения задачи SLAM для статических сцен, однако многие приложения

робототехники и автономной навигации требуют разработки методов для точного решения проблемы SLAM на динамических трехмерных сценах со множеством препятствий. Поэтому актуальным направлением развития методов SLAM является разработка методов слежения за объектами на динамических сценах. Успешное обнаружение объектов и их классификация объектов являются необходимыми условиями для решения задачи слежения за объектами на динамических сценах. Приведем несколько важных подходов к решению задачи слежения за объектом на динамических сценах. Классический подход к реализации поставленной задачи основан на применении фильтра Калмана для отслеживания положений объекта в пространстве. Фильтр Калмана предназначен для рекурсивной оценки вектора состояния известной динамической системы и использует как оценки вектора состояния, так и его оценки неопределенности на основе формулы условной вероятности Байеса. Однако лучшие результаты слежения с использованием фильтра Калмана достигаются посредством Гауссовского распределения входных данных. Идеи фильтрации Калмана, Марковских цепей и метода Монте-Карло активно используются для создания новых алгоритмов слежения за частицами. Перечисленные методы обладают высокой вычислительной сложностью и имеют проблемы со сходимостью. В исследовании рассмотрено решение задачи планирования траектории, при решении которой используется информации о результатах слежения за объектами на динамических сценах. В работе используются методы слежения за несколькими объектами (динамическими препятствиями) на основе бесконтактного распознавания объекта на расстоянии по визуальным данным и на основе методов фазового одновременного предсказания движения нескольких объектов по динамическим данным. Данный подход решения задачи является свободным от типичных проблем существующих методов слежения, таких как исчезновение объекта из поля зрения, частичное закрытие объекта другим объектом, одновременное слежение за несколькими объектами, неравномерное освещение, наличие иных пространственно-неоднородных искажений.

5.1 Решение задачи навигации мобильной платформы в неизвестном пространстве на основе расширенного фильтра Калмана и семантических маркеров

На основе предложенного в главе 2 метода регистрации данных была разработана методология решения задачи одновременной навигации и построения карты, основанная на комбинировании визуальных и семантических характеристик неизвестной окружающей среды. Данные семантические маркеры (навигационные ориентиры) используются при решении задачи определения динамического положения мобильной платформы, а также при решении задачи планирования траектории движения платформы из исходной точки в целевую точку в пространстве с препятствиями [80-88]. Предложенный метод определения динамического положения мобильной платформы на основе расширенной Калмановской фильтрации (EKF), истории положений камеры и перемещений платформы, семантических маркеров и совмещении полученных трехмерных карт глубины сочетает лучшие характеристики точных алгоритмов, основанных на вариациях Калмановской фильтрации, [51, 53] и быстрых алгоритмов, основанных на навигации по ориентирам [52].

EKF SLAM хорошо известный метод для решения задачи навигации в неизвестном пространстве. Основные шаги метода на основе EKF SLAM включают: прогнозирование траектории движения, инициализацию новых навигационных ориентиров и коррекцию расположения навигационных ориентиров на сцене (Рис. 5.1). Определим модель наблюдения для EKF SLAM

$$y = h(x) + V, \quad (5.1)$$

где x – вектор состояния мобильной платформы и ориентиров, $h(\cdot)$ – функция наблюдения и V – шум процесса.

Тогда ковариационная матрица для фильтра Калмана

$$P = FPF' + Q, \quad (5.2)$$

где F – якобиан, Q – Гауссовский шум.

Шаг коррекции метода EKF SLAM записывается как

$$z = y - h(x'), Z = HPx^T + R, \quad (5.3)$$

где z – обновленное значение положения, x' – положение мобильной платформы в модели наблюдения, H – якобиан, R – матрица ковариаций измеренных шумов процесса.

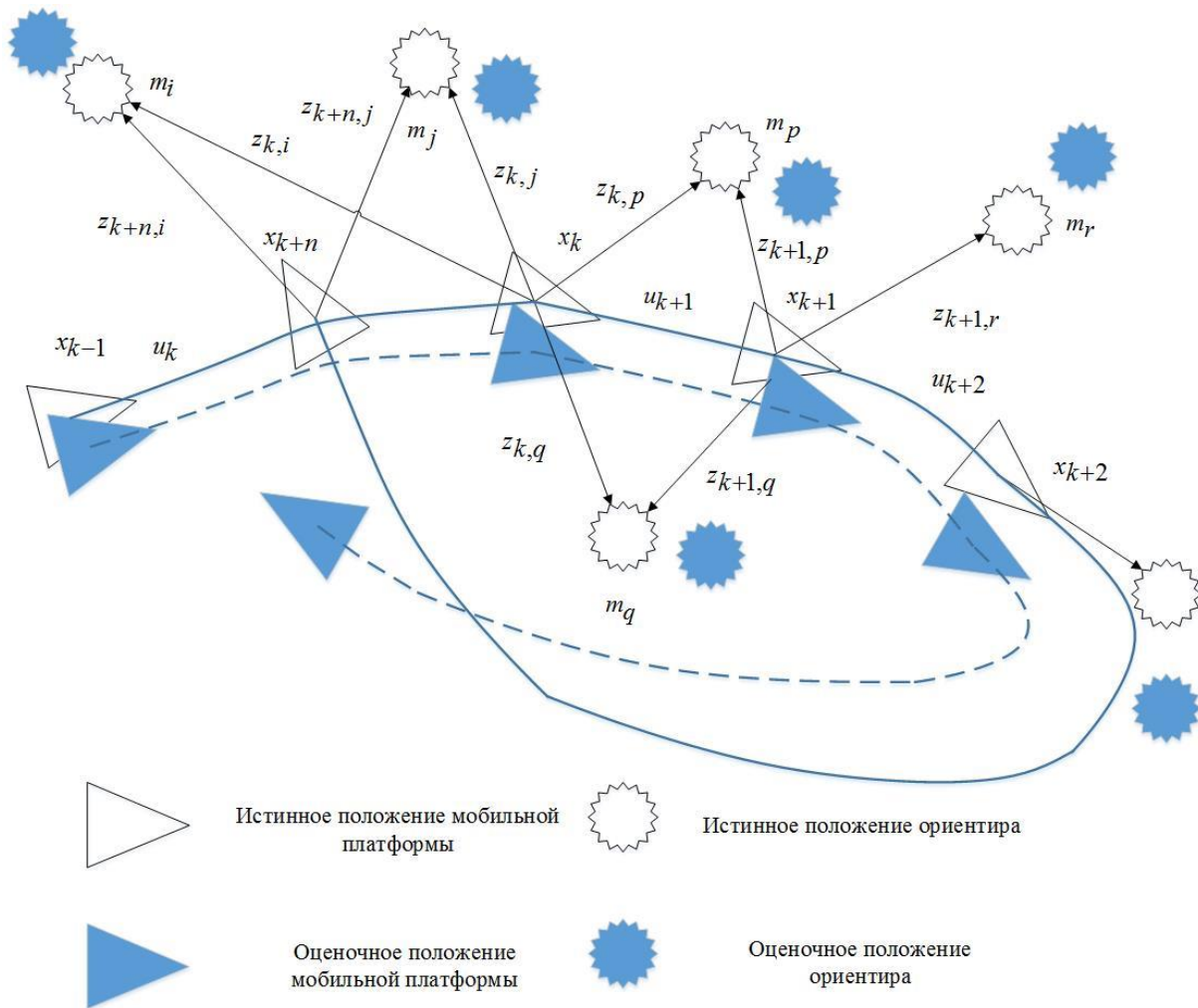


Рисунок 5.1 – Графическая интерпретация метода EKF SLAM для навигации мобильной платформы в пространстве

Коэффициент усиления (ядра) фильтра Калмана K

$$K = PH^T Z. \quad (5.4)$$

Обновленное положение мобильной платформы на трехмерной сцене может быть определено следующим образом

$$x'^p = x' + Kz, \quad (5.5)$$

где x'^p – значение предсказания положения мобильной платформы в текущий момент времени, x' – значение для положения мобильной платформы на сцене в прошлый момент времени.

Тогда получим ковариационную матрицу коррекции как

$$P^p = P - KZK^T. \quad (5.6)$$

где P^p – значение для предсказания ошибки, P – величина ошибки в прошлый момент времени. Данные измерения получены при помощи лазерных дальномеров, включая расстояние, угол между мобильной платформой и положением навигационных ориентиров на сцене.

Для предсказания положения мобильной платформы в исследовании разработан новый метод одновременной навигации и картографирования S-EKF SLAM [24, 32] (Semantic Extended Kalman Filter SLAM) на основе локализации и навигации с использованием семантических характеристик окружающей среды. Пусть k – положение мобильной платформы на сцене для соответствующего шага метода S-EKF SLAM. Тогда $x_v(k)$ – вектор состояния на шаге k ; $M = \{m_1, \dots, m_n\}$ – множество навигационных ориентиров на трехмерной сцене; n – текущее количество ориентиров.

Определим вектор состояния фильтра Калмана при перемещении мобильной платформы на сцене в виде линейного уравнения перехода с дискретным временем

$$x_v(k+1) = JT_v(k) \text{ dist}(x_v(k), S_v(k+1)) + U_v(k+1) + V_v(k+1), \quad (5.7)$$

«где dist – функция формирования вектора навигационных ориентиров, $U_v(k)$ – вектор управляющих входов, $JT_v(k)$ – матрица перехода состояния и $V_v(k)$ – вектор временно несогласованных ошибок шума процесса. В модель процесса расширенного фильтра Калмана добавлены семантические навигационные ориентиры $S_v(k+1)$, которые позволяют улучшить сходимость алгоритма определения динамического положения мобильной платформы за счет сокращения числа навигационных ориентиров и размерности вектора состояния» [8].

Предложенный метод решения проблемы одновременной навигации и картографирования S-EKF SLAM позволяет динамически получать точные и непрерывно обновляемые оценки положения мобильной платформы на трехмерной сцене с использованием дискретного ряда времени, содержащего неточные измерения положения платформы на сцене. В предлагаемом методе комбинирование информации происходит на уровне выделенных семантических и мультимодальных визуальных характеристик сцены.

В работе был апробирован известный подход к оптимизации графа положений мобильной платформы и обнаружению «замыканий цикла» на основе метода корзина слов (глава 7). Один из известных подходов к решению данной задачи был встроен в общую схему предлагаемого подхода к одновременной навигации и построения карты на основе S-EKF SLAM. В работе были разработаны эффективные параллельные алгоритмы определения динамического положения мобильной платформы в относительной системе координат на основе S-EKF SLAM. На рисунке 5.2 представлена обобщенная схема алгоритма S-EKF SLAM для навигации мобильной платформы в неизвестном пространстве.

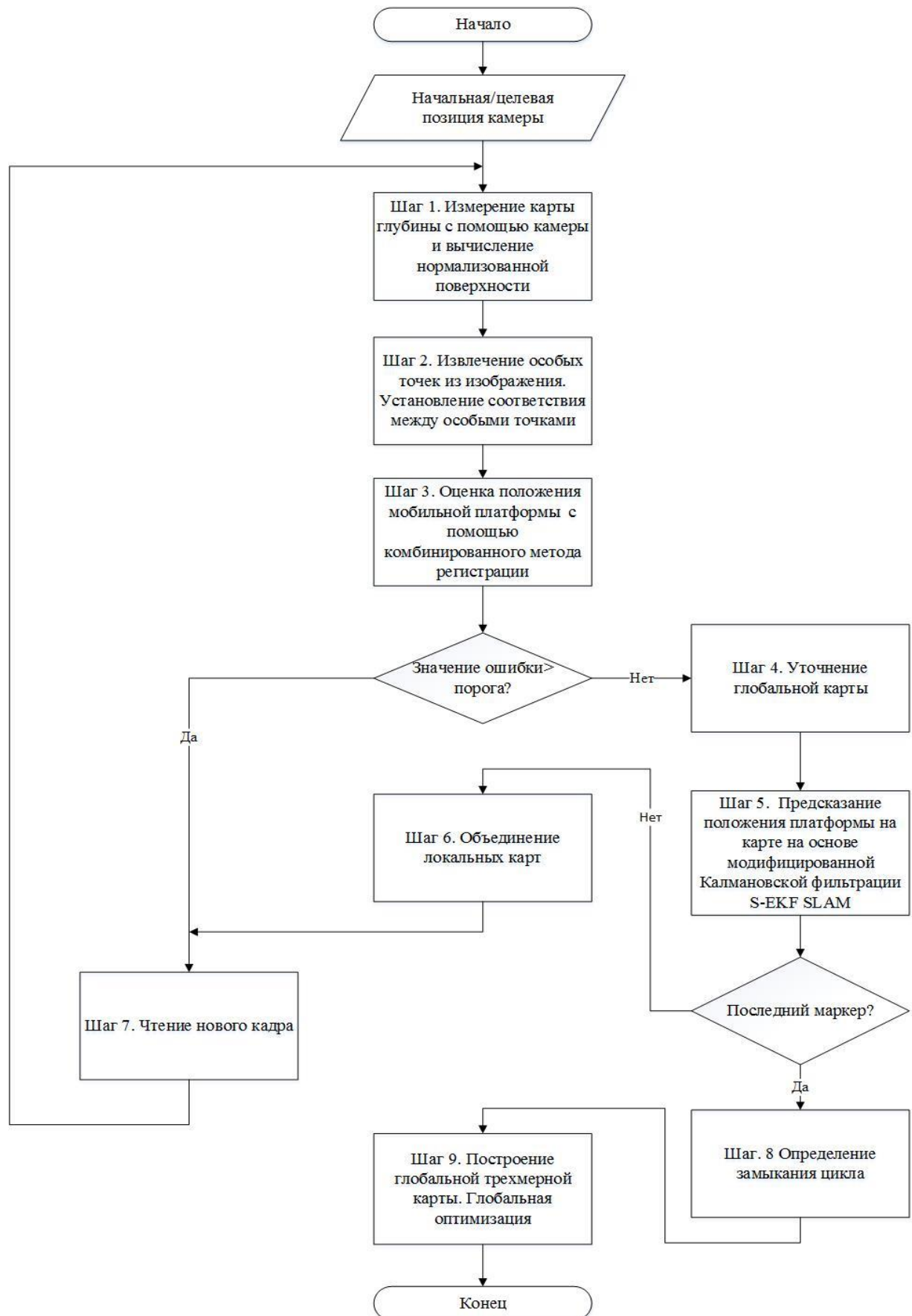


Рисунок 5.2 – Обобщенная блок-схема алгоритма S-EKF SLAM для навигации мобильной платформы в неизвестном пространстве

5.2 Идентификация семантических маркеров на сцене. Метод решения задачи локализации и навигации S-EKF SLAM

Для организации взаимодействия между человеком и мобильной платформой на сцене автономные интеллектуальные объекты должны понимать среду, в которой они работают. Современные алгоритмы SLAM основаны на использовании семантической сегментации и картографирования, классификации и распознавании трехмерных объектов на сцене. Объекты на сцене имеют свойства: семантические свойства (принадлежность к классу), визуальные свойства (данные о количестве и пространственном распределении особых точек в кадре), размеры, данные о подвижности объекта (статический или динамический), данные о траектории перемещения для динамических объектов, структурные характеристики для включения их отличительных признаков; набор ссылок, указывающих на принадлежность к исследуемой сцене.

Информация о семантических свойствах объектов на сцене используется при решении следующих задач:

1. Вариационная задача точка-плоскость для класса ортогональных преобразований: принадлежность к определенному классу предметной области влияет на значение весовых коэффициентов, которые связаны с семантической маркировкой окружающего пространства.
2. Вариационная задача точка-точка для класса аффинных преобразований: принадлежность к определенному классу предметной области влияет на значение весовых коэффициентов, которые связаны с семантической маркировкой окружающего пространства.
3. Определение положения мобильной платформы в неизвестном пространстве: семантические маркеры учитываются при расчете матрицы перехода состояний и вектора управлений.

4. Планирование траектории мобильной платформы в неизвестном пространстве: информация о семантической маркировке пространства учитывается при разметке метрико-топологического графа.

Для определения семантических свойств объектов на сцене используется эталонная база данных NYU Depth Dataset v2 [179], которая содержит размеченный набор данных RGB-D кадров с сегментацией.

В приложениях с мобильными платформами методы семантической сегментации могут быть усовершенствованы путем наблюдения за окружающей средой с нескольких точек зрения. В современных системах автономной навигации для семантической маркировки пространства используется конвейер методов, которые решают следующие задачи:

- идентификация объектов (определение всех пикселей объектов одного или нескольких классов на сцене и разметка сцены при помощи рамки, которая определяет границы для каждого объекта и содержит метку с вероятностью отнесения объекта к определенному классу), например, детектор YOLO [183, 184];
- семантическая сегментация (определение всех пикселей объектов определённого класса или фона на изображении), например, метод 3DEF [185];
- классовая сегментация (определение всех пикселей объекта определённого класса по отдельности для каждого объекта класса на сцене), например, детектор Mask R-CNN;
- сегментация изображений, основанная на разрезах графа, например, метод GrabCut [188].

Сверточная нейронная сеть Mask R-CNN развивает архитектуру сети Faster R-CNN путём добавления ещё одной ветки анализа, которая предсказывает положение маски, покрывающей найденный объект. Детектор YOLO выделяет только рамки объектов, а детектор Mask R-CNN выделяет рамки и считает вероятность обнаружения или промаха для каждого выделенного объекта.

Информация о семантических характеристиках окружающей среды в диссертационной работе получена в результате использования известных методов семантической сегментации окружающего пространства: YOLO [183, 184] и 3DEF

[185], которые позволяют решать известную в робототехнике проблему привязки символов. Комбинация различных методов семантической обработки RGB-D данных позволяет осуществлять точную классификацию объектов разных классов. Известно, что методы YOLO, Mask-RCNN, 3DEF дают разные оценки по точности для разных классов объектов – структурные элементы сцены, пол, потолок, стены; окна, плоские объекты, такие как картины, рамки, книги; выпуклые ригидные объекты, такие как стол, стул, софа, и.т.д.

В результате постоянно повторяющейся последовательности шагов при решении проблемы SLAM формируется множество абстрактных состояний и свойств окружающей среды, которые постоянно пополняют концептуальные знания мобильной платформы об окружающем мире. В работе для представления семантики использовано символьное представление для множества понятий. В работе представлен метод идентификации и представления свойств неизвестного окружающего пространства на трехмерной сцене в семантической форме [29]. В данной работе используется детектор объектов YOLO (You Only Look Once) [183, 184] и метод трехмерной семантической сегментации на основе 3D-запутанных лесов (3D Entangled Forest, 3DEF) [185] для определения принадлежности объектов к классам предметной области. На рисунке 5.3 представлен пример семантической маркировки помещения с использованием детектора объектов (YOLO) и метода трехмерной семантической сегментации сцены на основе 3DEF для реального вычислительного эксперимента.

В предлагаемом подходе к семантической сегментации сцены трехмерное облако точек разбивается на супервоксели с помощью метода сегментации Voxel Cloud Connectivity (VCCS) [186]. На следующем этапе этот подход применяет алгоритм роста области, который рекурсивно объединяет два смежных сегмента на трехмерной сцене s_i и s_j в более крупные сегменты. Этот шаг слияния выполняется на основе оценки функции расстояния $\text{dist}(s_i, s_j)$, которая представляет собой линейную комбинацию цвета, нормали поверхности и расстояния от точки до плоскости между сегментами. Для каждого полученного сегмента s_i вычисляется вектор пространственных объектов X длиной 18.

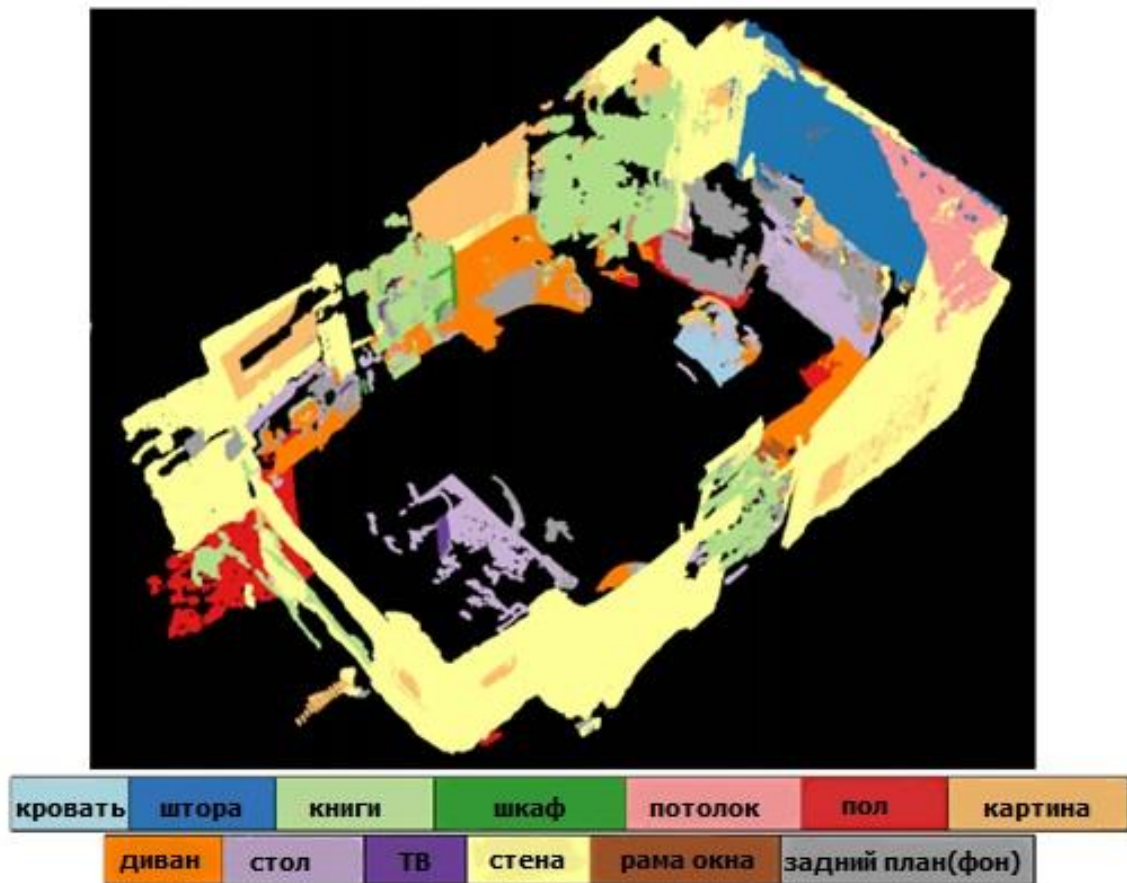


Рисунок 5.3 – Пример семантической маркировки помещения с использованием детектора объектов YOLO и метода трехмерной семантической сегментации сцены 3DEF

Помимо простых цветовых функций, вектор включает в себя быстрые геометрические функции. Затем выполняется процедура семантической классификации: для каждого сегмента s_i выбирается набор близких сегментов s_j на основе трех ограничений: расстояние от точки до плоскости, замкнутые углы и евклидово расстояние. Во время процедуры обучения и вывода данный набор сегментов используется для оценки набора двоичных тестов, определяющих запутанные объекты, которые способны описывать сложные геометрические отношения между сегментами в окрестности.

Недостатком данного классификатора 3DEF является зависимость точности классификации от позиции обзора мобильной платформы, для устранения этого недостатка в работе используется несколько представлений сцены (Multi-viewed

3D Entangled Forest) [187]. Семантический метод трехмерной сегментации, такой как 3DEF, может точно сегментировать многие структурные элементы сцены такие как пол, потолок, стена, кровать, диван, стул. Данный метод плохо классифицирует сложные классы объектов, такие как окно, телевизор, картины, различные малые формы интерьера, расположенные/установленные в шкафах, так как их семантическая сегментация основана информации о цвете. Для устранения указанного недостатка в работе при классификации данных объектов используется детектор объектов YOLOv3 [184].

Для интеграции детектора объектов YOLOv3 в предложенный подход необходимо: выделить объект в кадре при помощи детектора; внутри ограничительной рамки детектора провести процедуру сегментации объекта вдоль его контуров на основе метода Grabcut [188]. GrabCut – это метод сегментации изображений, основанный на разрезах графа, в котором взвешенный граф строится на основе соседних пикселей и помеченных масок. Далее создается сегментация фона или переднего плана, сочетающая жесткую сегментацию путем итеративной оптимизации графа с матированием границ, для того чтобы работать с размытыми и смешанными пикселями на границах объекта. Детектор Grabcut используется как для данных о цвете, так и для данных о глубине RGB-D кадра. Сегментации, полученные из RGB данных и данных о глубине, можно комбинировать с помощью операции логическое ИЛИ. Для каждого обнаруженного объекта метод Grabcut инициализируется ограничительной рамкой детектора YOLOv3 для надежности, так как Grabcut не всегда может найти сегментацию, количество итераций в методе равно 5. Если метод Grabcut не может найти сегментацию, то существует две альтернативы: пропуск обнаружения или рассмотрение всего ограничивающего поля в качестве переднего плана. После этого шага для каждого сегментированного пикселя в рамке можно сохранить метки YOLOv3 и метки Grabcut, при этом ограничительные рамки детектора и метода Grabcut в кадре могут перекрываться, поэтому порядок обработки ограничительных рамок может отрицательно повлиять на результаты работы конвейера методов в целом. Для надежной семантической сегментации ограничительные рамки YOLOv3 сортируются в порядке уменьшения

размера. Например, большие рамки могут быть опорными поверхностями, такими как столы, в то время как маленькие рамки могут содержать объекты интерьера, например, книги [35]. Объединение результатов семантической сегментации метода 3DEF и комбинации методов YOLOv3 + Grabcut осуществляется с использованием слияния на основе Байесовского подхода. Представим сказанное выше в виде двух алгоритмов: алгоритм 1 – алгоритм семантической сегментации на основе уточнения положения детектирующей рамки; алгоритм 2 – алгоритм семантической сегментации, комбинирующий 3DEF алгоритм сегментации и алгоритм сегментации на основе уточненных ограничительных рамок.

Алгоритм 1. Алгоритм семантической сегментации на основе уточнения положения детектирующей рамки.

Входные данные: последовательность кадров в виде пары изображений и карт глубины (I-RGB).

Выходные данные: сегментированное изображение, размеченное метками и вероятностями.

Шаг 1. Детектирование объектов на изображении I-RGB при помощи метода YOLOv3 – определение ограничительных рамок у изображений.

Шаг 2. Сортировка списка ограничительных рамок (priors, anhcors) в порядке уменьшения размера площади рамок. Данный этап решает проблему перекрытия рамок во время процедуры детектирования.

Шаг 3. Выполнить в цикле для каждой ограничительной рамки в списке следующие действия:

- выполнить семантическую сегментацию данных на изображении I-RGB при помощи метода Grabcut;
- выполнить семантическую сегментацию данных на карте глубины I-D при помощи метода Grabcut. Выполним метод оптимизации разреза графа для 5 итераций;
- для надежности сегментации, данные полученные из кадров RGB и глубины, объединяются с помощью операции логическое ИЛИ пиксель на пиксель;

– проверяем условие: если метод Grabcut не может вернуть какую-либо сегментацию, мы рассматриваем начальные рамки обнаруженного объекта как передний план, иначе ограничительная рамка попадает в список ограничительных рамок, которые подлежат разметке.

Шаг 4. Разметка изображения на основе списка ограничительных рамок, полученного на Шаге 3. Для каждого элемента этого списка выбирается соответствующее ему изображение и каждому пикселю данного изображения присваивается метка с названием класса объекта и степень вероятности отнесения к данному классу.

Алгоритм 2. Комбинированный алгоритм семантической сегментации с использованием алгоритма 1 и метода 3DEF.

Входные данные: кадр, содержащий 2D и 3D данные – изображение, карта глубины, трехмерное облако точек.

Выходные данные: сегментированное размеченное сглаженное 3D облако точек с метками и значениями вероятностей.

Шаг 1. Выполнить Алгоритм 1 для семантической сегментации с использованием информации о цвете и карте глубины. Получить размеченное изображение с метками классов и значением вероятности отнесения сегментированного объекта к классу.

Шаг 2. Применить метод семантической сегментации 3DEF к данным в виде трехмерного облака точек. На выходе метода получаем список классов и размеченное облако точек. Метод 3DEF состоит из следующих шагов: сверхсегментация супервокселей в 3D-патчах; объединение похожих соседних сегментов в более крупные, преимущественно плоские сегменты; классификация сегментов.

Шаг 3. Для каждого кадра выполнить слияние выходных данных, полученных на Шаге 1 (размеченное изображение) и Шаге 2 (размеченное облако точек) на основе Байесовского подхода. Слияние применяется к значениям вероятности принадлежности объектов к классу. На выходе данного шага получаем новое размеченное трехмерное облако точек.

Шаг 4. Сглаживание данных. Этот шаг учитывает контекст пикселя, чтобы повысить надежность в отношении ошибок в процессе прямой проекции, которые могут быть вызваны шумом или локально неточной регистрацией. Каждая точка в данных вносит свой вклад не только своей меткой, но и учитывает значение самой частой метки в ее окрестности Мура. На выходе получаем размеченное сглаженное 3D облако точек.

На рисунке 5.4 представлен пример автоматической семантической маркировки помещения с использованием эталонной базы данных NYU Depth Dataset (набор данных Classrooms).



Рисунок 5.4 – Пример семантической маркировки помещения с использованием детектора объектов (YOLOv3) и метода семантической сегментации сцены на основе 3DEF

Черным и белым цветом на рисунке отмечены структурные элементы помещения: стены, пол, потолок и окна. Остальные классы – столы, стулья, шкафы, книги, ковровое покрытие, учебная доска отмечены оттенками красного в зависимости от значений w_i, w_j - весовых характеристик данных, связанные с семантической маркировкой пространства.

Предложенный подход к решению проблемы SLAM (S-EKF SLAM) [24] базируется на использовании мультисенсорной динамической информации, на пространственном согласовании последовательных кадров данных, на локализации семантических маркеров, на слежении за объектами на трехмерной сцене, на обнаружении «замыканий цикла» (возвращение мобильной платформы к месту,

которое было посещено ранее), на согласовании полной последовательности кадров и глобальной оптимизации (происходит одновременное согласование и регистрация последовательности изображений и соответствующих трехмерных облаков точек).

Определим требования к входным данным и объектам сцены:

1. Мобильная платформа определяет свое положение в пространстве относительно целевой точки.
2. Объекты на сцене имеют жесткую форму (ригидные).
3. Степень свободы движения мобильной платформы 6D.
4. При планировании траектории необходимо учитывать конструктивные и динамические ограничения мобильной платформы.
5. Данные сцены получены на основе следующих датчиков: камера глубины, камера видимого диапазона, времяпролетная камера.
6. Одометрическая информация не используется.
7. План окружающей среды для мобильной платформы не существует.
8. Среда содержит препятствия, в том числе динамические препятствия; для проведения экспериментов предлагается рассматривать в качестве препятствий объекты определенного размера на сцене.
9. Точность оценки траектории движения мобильной платформы определяется типом используемых датчиков, комплексом семантических маркеров и временными ограничениями.
10. Движение мобильной платформы к целевой точке должно быть оптимальным по времени.

Детализированная схема предлагаемого метода одновременной навигации и построения карты на основе S-EKF SLAM вокруг мобильной платформы в окружающей среде представлена на рисунке 5.5.

Трехмерная карта сцены представляет собой вектор состояния большой размерности, который содержит положения мобильной платформы и навигационных ориентиров:

$$x = \begin{bmatrix} R \\ M \end{bmatrix} = \begin{bmatrix} R \\ L_1 \\ \vdots \\ L_N \end{bmatrix}, \quad (5.8)$$

где R – положение мобильной 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 $\bar{x}$ и C :

$$\bar{x} = \begin{bmatrix} \bar{R} \\ \bar{M} \end{bmatrix} = \begin{bmatrix} \bar{R} \\ \bar{L}_1 \\ \vdots \\ \bar{L}_N \end{bmatrix}, \quad \begin{bmatrix} C_{RR} & C_{RM} \\ C_{MR} & C_{MM} \end{bmatrix} = \begin{bmatrix} C_{RR} & C_{RL_1} & \dots & C_{RL_N} \\ C_{L_1R} & C_{L_1L_1} & \dots & C_{L_1L_n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{L_NR} & C_{L_nL_1} & \dots & C_{L_nL_n} \end{bmatrix}. \quad (5.9)$$

Таким образом, целью EKF SLAM является постоянное обновление карты $\{\bar{x}, C\}$. В классическом методе EKF, если x – вектор состояния, u – вектор управления и n – вектор возмущения, можно записать обобщенную временную функцию обновления $x \leftarrow f(x, u, n)$.

Шаг предсказания в классическом алгоритме EKF $\{x, C\}$ может быть записан следующим образом

$$\bar{x} \leftarrow f(\bar{x}, u, 0), \quad C \leftarrow F_x C F_x^T + F_n N F_n^T. \quad (5.10)$$

В EKF SLAM только часть состояния является вариантом времени: мобильная платформа, которая движется. Поэтому имеет место различное поведение для каждой части вектора состояния. Тогда поскольку большая часть трехмерной карты сцены инвариантна относительно движения мобильной платформы, то можно представить разреженные матрицы якоби со следующей структурой:

$$F_x = \begin{bmatrix} \frac{dF_R}{dR} & 0 \\ 0 & I \end{bmatrix}, \quad F_y = \begin{bmatrix} \frac{dF_R}{dn} \\ 0 \end{bmatrix}. \quad (5.11)$$

Тогда можно представить разреженные уравнения для прогнозирования траектории движения мобильной платформы в трехмерном пространстве с помощью расширенного фильтра Калмана в следующем виде:

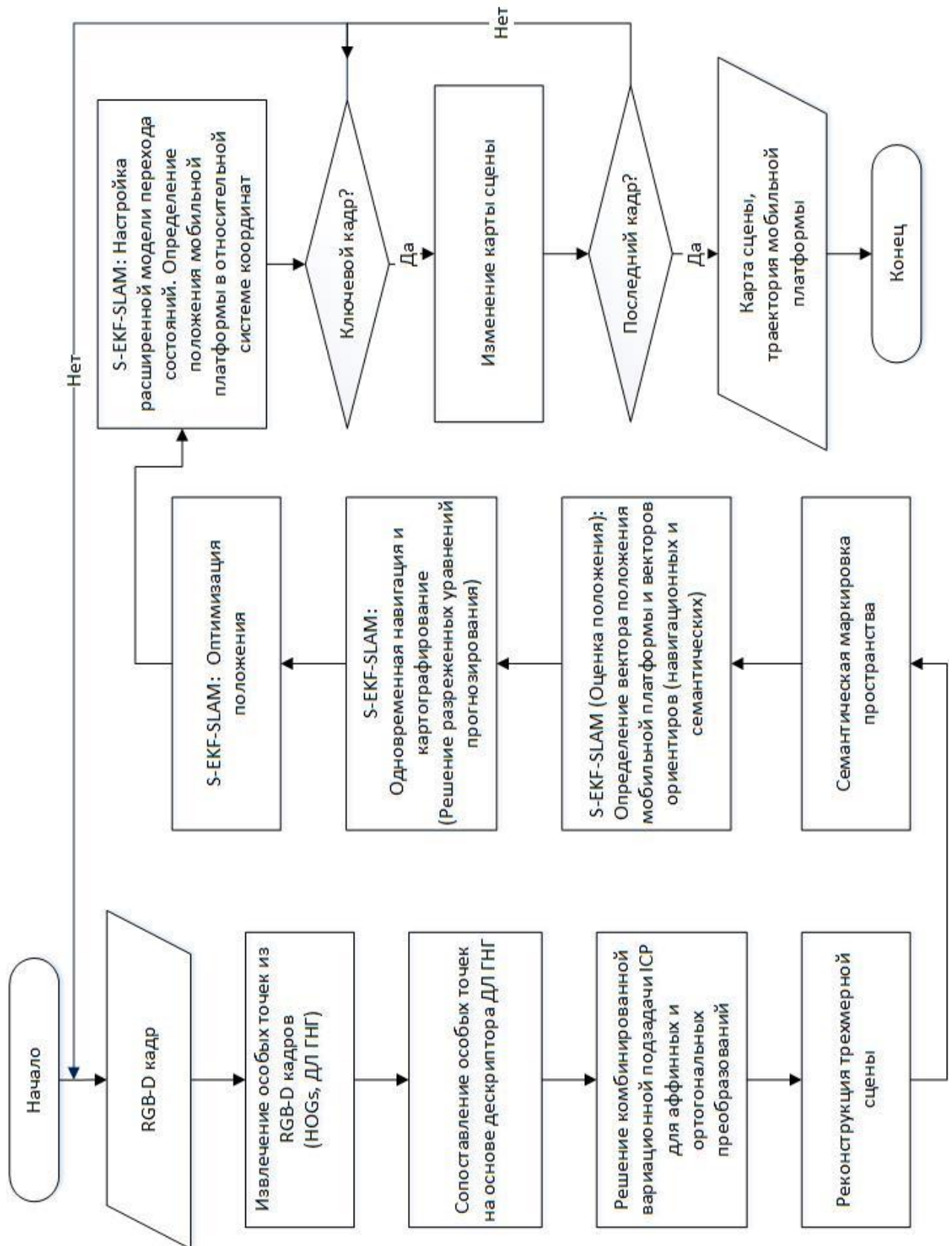


Рисунок 5.5 – Детализированная блок-схема алгоритма навигации мобильной платформы в неизвестном пространстве на основе S-EKF SLAM

$$\begin{aligned}
\bar{R} &\leftarrow f_R(\bar{R}, u, 0), \\
C_{RR} &\leftarrow \frac{dF_R}{dR} C_{RR} \frac{dF_R^T}{dR} + \frac{dF_R}{dn} N \frac{dF_R^T}{dn}, \\
C_{RM} &\leftarrow \frac{dF_R}{dR} C_{RM}, \\
C_{MR} &\leftarrow C_{RM}.
\end{aligned} \tag{5.12}$$

Алгоритмическая сложность этого набора уравнений составляет $O(n)$. Мобильная платформа оснащена мультисенсорными камерами, которые измеряют относительное расположение между любым семантическим ориентиром и самой мобильной платформой. Абсолютное расположение семантических ориентиров (маркеров) заранее неизвестно. Реализация предложенного метода S-EKF SLAM использует нелинейные кинетические модели мобильной платформы и нелинейные асинхронные модели наблюдения в контролируемых/неконтролируемых условиях. Состояние исследуемой системы в каждый момент времени состоит из данных о положении и ориентации мобильной платформы, а также данных о положении всех семантических маркеров относительно платформы. Состояние мобильной платформы на временном шаге k обозначается $x_v(k)$. Движение платформы через окружающее пространство моделируется обычным линейным уравнением перехода дискретно-временного состояния или моделью процесса вида

$$p_i(k+1) = p_i(k) = p_i. \tag{5.13}$$

Количество навигационных визуальных ориентиров на сцене задается набором p с размерностью M и обозначается как вектор визуальных ориентиров:

$$p = [p_1^T \dots p_M^T]^T. \tag{5.14}$$

Функцию перехода состояния для i -го семантического ориентира можно представить следующим образом

$$s_j(k+1) = s_j(k) = s_j. \tag{5.15}$$

Вектор всех N семантических ориентиров может быть представлен

$$s = [s_1^T \dots s_N^T]^T. \tag{5.16}$$

Тогда получим расширенный вектор состояния, содержащий как состояние визуальных маркеров, так и состояние всех семантических маркеров:

$$x(k) = [x_v^T(k)p_1^T \dots p_M^T, s_1^T \dots s_N^T]^T = [x_v^T(k)ps_1^T \dots ps_N^T]^T. \quad (5.17)$$

Таким образом, расширенная модель перехода состояния для всей системы с учетом 5.17 может быть записана как

$$\begin{aligned} \begin{bmatrix} x_v(k+1) \\ p_1 \\ \vdots \\ p_N \end{bmatrix} &= \begin{bmatrix} JT_v(k+1) & 0 & \dots & 0 \\ 0 & L_{sp_1} & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & L_{sp_N} \end{bmatrix} \text{dist} \left(\begin{bmatrix} x_v(k) \\ p_1 \\ \vdots \\ p_M \end{bmatrix}, \begin{bmatrix} x_v(k) \\ s_1 \\ \vdots \\ s_N \end{bmatrix} \right) + \\ &+ \begin{bmatrix} U_v(k+1) \\ O_{p_1} \\ \vdots \\ O_{p_N} \end{bmatrix} + \begin{bmatrix} V_v(k+1) \\ O_{p_1} \\ \vdots \\ O_{p_N} \end{bmatrix}, \end{aligned} \quad (5.18)$$

где L_{sp_i} есть $\dim(p_i) \times \dim(p_i)$ – единичная матрица и O_{p_i} 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 . Так как число семантических ориентиров N в общем случае не совпадает с числом навигационных ориентиров M , последних на сцене значительно больше и всегда справедливо утверждение « M больше или равно N », то искусственным образом сузим вектор семантических ориентиров p до размерности N . Использование семантических ориентиров позволяет сократить размерность вектора состояния при решении проблемы локализации. Обновление истории положений и перемещений мобильной платформы, а также настройка исследуемых характеристик осуществляются с определенным интервалом времени. Новые данные о положении платформы на сцене добавляются не на каждой итерации фильтра Калмана, а в случае, если положение мобильной платформы существенно изменилось, например, при значительном смещении или вращении по сравнению с предыдущим положением. Математическая модель сцены сохраняется в виде графа, вершины которого соответствуют определенным моментам времени.

5.3 Компьютерное моделирование

В данном параграфе представлены и обсуждены результаты компьютерного моделирования. Для проведения моделирования применялся компьютер на базе Intel Core i7 с графическим процессором. Оценим точность и сходимость предложенного метода на примере эталонной базы данных NYU Depth Dataset V2 [179]. Для этого выберем следующие тестовые наборы данных: Classrooms (классная комната), Living Rooms (1/4, жилая комната, фрагмент сцены, полученный с камеры с узким углом обзора), Offices (1/2, жилая комната, фрагмент сцены, полученный с камеры с широким углом обзора), Offices (2/2, офис №2, фрагмент сцены, полученный с камеры с широким углом обзора).

Набор данных NYU Depth Dataset состоит из нескольких компонентов:

1. Размеченный набор данных: состоит из выравненных RGB-D кадров (изображения и данные глубины), которые были синхронизированы и снабжены аннотациями с метками семантической сегментации для каждого изображения.
2. Необработанный набор данных: содержит raw данные изображений, глубины и акселерометра, которые были получены с использованием камеры Kinect 2.0.
3. Toolbox: интерфейс данных, который содержит функции для доступа и управления RGB-D данными и метками.

Каждый набор данных содержит файлы дампа, изображение в формате ppm (Рис. 5.6, а), данные о глубине в формате pgm (Рис. 5.6, б), данные о семантической сегментации сцены (Рис. 5.6, в). Размеченный набор данных NYU Depth Dataset содержит 13 классов объектов, связанных с основными объектами в помещении: структурные элементы сцены, окна, двери, мебель, техника. Результаты семантической сегментации размеченного набора данных использованы как в качестве обучающей выборки для нейронной сети, так и в качестве эталонного набора данных, на котором осуществляется проверка точности

предлагаемого метода регистрации данных и известных методов трехмерной реконструкции сцены.

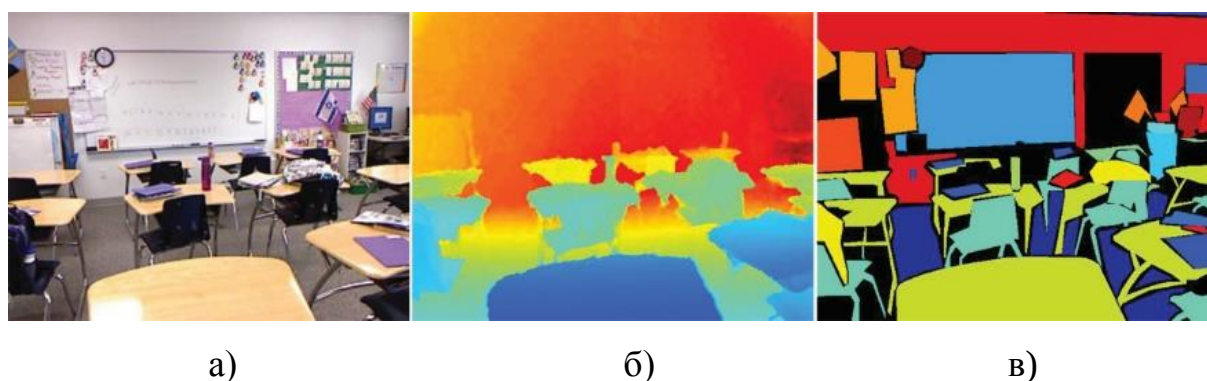


Рисунок 5.6 – Тестовый набор данных классная комната (Classrooms) из NYU Depth Dataset: а) визуальные данные кадра, б) данные глубины кадра, в) данные о семантической сегментации кадра

Аналогично другим известным подходам, которые используют данный набор для оценки эффективности, в работе используется три показателя: общая точность (ОТ), классовая полнота (КП), классовая точность (КТ). Последние два показателя менее зависимы от значения частоты классов на анализируемой сцене в целом. Проанализируем различные комбинации метода трехмерной сегментации 3DEF и детектора объектов YOLOv3 и оценим влияние использования нескольких представлений на точность метода семантической сегментации. Таблица 5.1 показывает, что интеграция детектора объектов YOLOv3 всегда значительно улучшает точность сегментации по сравнению с базовым методом 3DEF или его модификацией MV-3DEF. Метод на основе MV-3DEF+YOLO работает немного лучше, чем 3DEF + YOLO, но с другой стороны, первый метод требует существенно больших вычислительных ресурсов при слиянии данных о разных представлениях сцены, чем второй метод семантической маркировки пространства.

В таблице 5.2 и таблице 5.3 представлены результаты семантической маркировки пространства для набора данных Living Rooms (1/4) из базы NYU Depth Dataset, которые содержат значения классовой точности для структурных элементов сцены и для элементов обстановки помещения соответственно.

Таблица 5.1 – Точность и полнота для различных методов семантической маркировки окружающего пространства (в %)

Метод семантической маркировки	Классовая точность	Общая точность	Классовая полнота
3D Entangled Forest	54.8	65.7	55.5
Многовидовый 3DEF	57.2	65.9	54
3DEF и YOLOv3	63.4	68.3	56.9
Многовидовый 3DEF и YOLOv3	64.1	68.5	57.1

Таблица 5.2 – Точность (в %) методов семантической маркировки окружающего пространства для структурных элементов сцены в NYU Depth Dataset

Метод семантической маркировки	Наименование класса объекта					
	Стена	Пол	Окно	Потолок	Мебель	Дверь
3D Entangled Forest	85.1	98.5	25.8	80.5	49	46.1
Многовидовый 3DEF	85.3	98.5	30.1	80	49.7	50.9
3DEF и YOLOv3	83.6	97	25.1	79.1	59.1	52.8
Многовидовый 3DEF и YOLOv3	83.6	96.8	26.2	81.4	56.3	56.1

Как видно, в обеих таблицах значительное повышение классовой точности достигается за счет объединения классификатора 3DEF и детектора объектов YOLO для большинства классов объектов, которые представлены в наборе данных NYU Depth Dataset. Известно, что методы YOLO, Mask-R-CNN, 3DEF дают разные оценки по точности для разных классов – структурные элементы сцены, такие как пол, потолок, стены; окна, плоские объекты, такие как картины, рамки, книги; выпуклые ригидные объекты, такие как стол, стул, софа, и.т.д.

Таблица 5.3 – Точность (в %) методов семантической маркировки окружающего пространства для элементов обстановки помещения в NYU Depth Dataset

Метод семантической маркировки	Наименование класса объекта					
	Диван	Кресло	Картина	Стол	Книга	Телевизор
3D Entangled Fores	71.3	65.9	28.6	45.4	55.1	36.1
Многовидовый 3DEF	74.9	66.5	29.2	50.1	55.8	43.6
3DEF и YOLOv3	87.9	83.4	24.7	51.8	71.3	59.5
Многовидовый 3DEF и YOLOv3	88.2	83.1	23.2	50.9	72.1	58.2

Из таблицы 5.3 видно, что лучшие по точности результаты 3DEF+YOLO превосходят результаты на основе 3DEF или MV-3DEF (+5,2% в КТ, +2,3% в ОТ и +2,5% в КП), с другой стороны для таких классов как окно, картина, стена предложенная комбинация методов не дает существенного увеличения точности классификации, например, для класса окно получены следующие результаты (+5,2% в КТ, +2,3% в ОТ и +2,5% в КП).

Данное обстоятельство связано с тем, что метод трехмерной сегментации 3DEF плохо классифицирует объекты похожие на плоские коробки: при расчете показателя классовой точности данный метод получает штраф за обнаружение на сцене объектов, принадлежащих классам с большим количеством образцов в эталонном наборе данных. Данный недостаток 3DEF не может быть компенсирован детектором объектов YOLOv3, так как он не обучен на этих классах. Решение данной проблемы может быть основано на обучении детектора объектов для данных классов или с помощью модификации процедуры предварительной сегментации растущей области в 3DEF.

Результаты семантической маркировки пространства (информация о принадлежности объектов к классам на сцене) используются при решении комбинированной вариационной задачи метода ICP для группы ортогональных и для группы аффинных преобразований: принадлежность к определенному классу

предметной области влияет на значение весовых коэффициентов, которые связаны с семантической маркировкой окружающего пространства.

Все реальные данные в исследовании получены для помещения – коридор Челябинского государственного университета. Для исследования точности и сходимости предложенного комбинированного метода регистрации данных было проведено два эксперимента:

- в первом случае человек управлял самостоятельно мобильной платформой с установленными на ней камерой Beward B2720 и/или камерой Kinect 2.0 и двигался вдоль коридора университета;
- во втором случае камера MESA SR 4000 и/или камера Kinect 2.0 были установлены на мобильные платформы Husky UGV и/или Odyssey 6 Robotics, при этом платформы самостоятельно передвигались в коридоре при помощи метода EKF SLAM.

Платформы с аналогами систем (программное обеспечение) для тестирования были взяты с сайта openslam.org. Системы были протестированы в различных условиях: реальные вычислительные эксперименты были проведены при равномерном и неравномерном освещении, количество объектов на сцене и топологическая структура сцены менялась в разных сериях экспериментов, также менялось взаимное расположение и количество семантических маркеров на сцене.

На рисунке 5.7 и рисунке 5.8 представлены результаты второго вычислительного эксперимента: предложенный метод для локализации и навигации S-EKF SLAM был сравнен с известными методами решения проблемы SLAM в контролируемых условиях (с равномерным освещением и отсутствием различных шумов) и в неконтролируемых (с неравномерным освещением) соответственно.

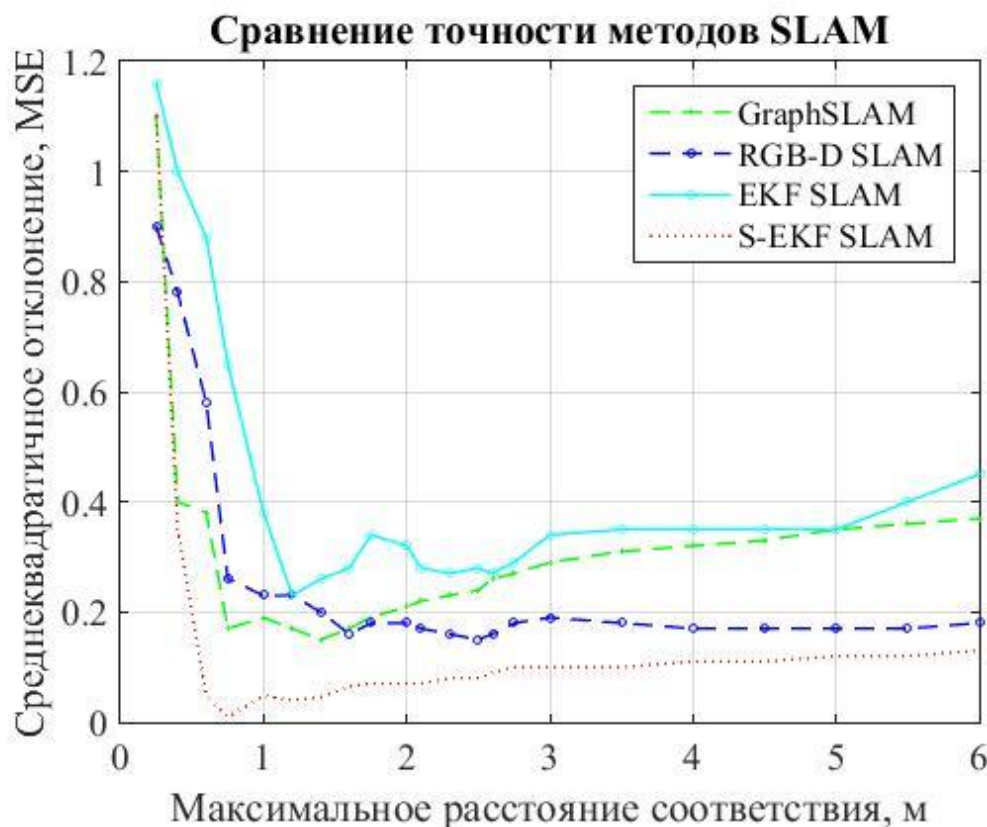


Рисунок 5.7 – Результаты вычислительного эксперимента: сравнительный анализ точности метода S-EKF SLAM и известных методов SLAM в контролируемых условиях (равномерное освещение)

В исследовании установлено: метод S-EKF SLAM при равномерном освещении для трехмерных сцен показывает похожие по точности результаты с методами на основе методологии Visual SLAM в терминах среднееквadraticного отклонения (MSE), при этом метод S-EKF SLAM имеет лучшую сходимость.

При неравномерном освещении для трехмерных сцен со сложной топологической структурой объектов метод S-EKF SLAM показывает лучшие по точности результаты в терминах среднееквadraticного отклонения (MSE), чем методы на основе методологий Visual SLAM, EKF SLAM, Graph SLAM, при этом сходимость метода S-EKF SLAM значительно улучшается при использовании процедуры RANSAC, а сходимость метода определения динамического положения мобильной платформы улучшается за счет использования семантических ориентиров в расширенном фильтре Калмана.

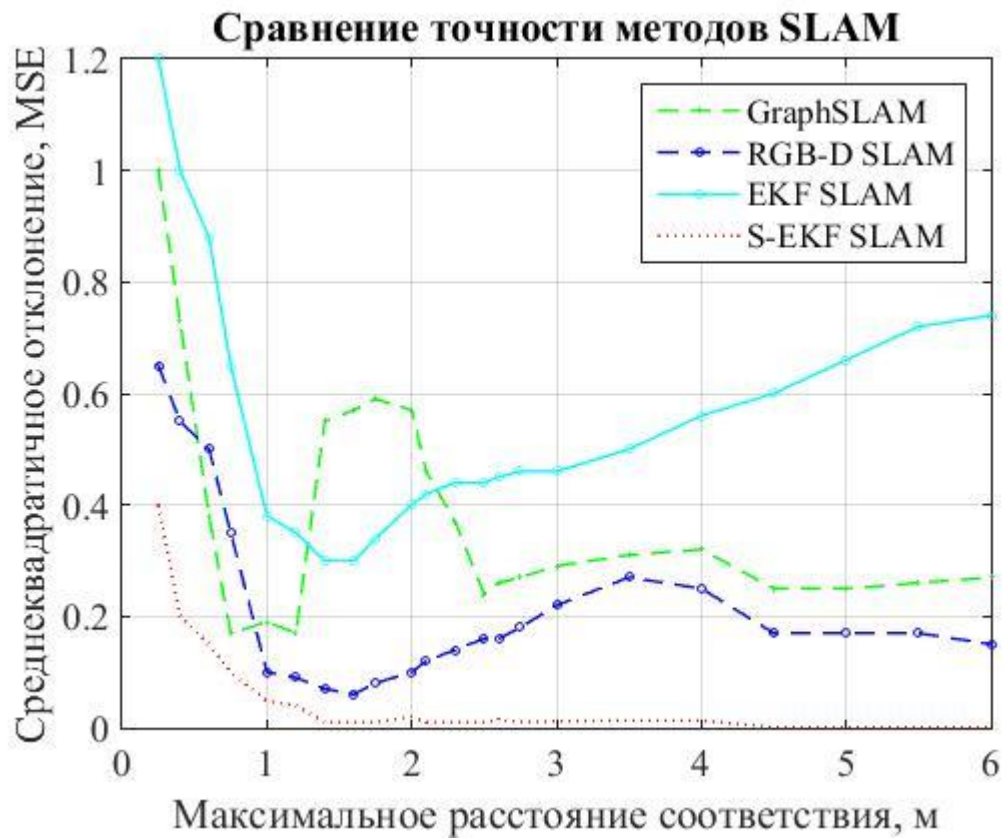


Рисунок 5.8 – Результаты вычислительного эксперимента: сравнительный анализ точности метода S-EKF SLAM и известных методов SLAM в неконтролируемых условиях (неравномерное освещение)

5.4 Алгоритм планирования траектории мобильной платформы в неизвестном пространстве на основе метрико-топологического графа и семантических маркеров

Во многих приложениях робототехники мобильные платформы, беспилотные летательные аппараты должны идентифицировать и обходить препятствия (непроходимые участки) в трехмерной среде. Для решения данной задачи автономные объекты должны создавать трехмерную карту динамической среды и определять на этой карте статические и динамические препятствия. Исследуемое пространство может содержать большое количество препятствий, не позволяющих построить близкую к оптимальной траекторию движения мобильной платформы на сцене для заданных начальной и целевой точек. Любой объект на сцене можно рассматривать как семантический ориентир, препятствие или фон. С точки зрения работы алгоритма навигации мобильной платформы для построения карты и одновременной локализации нас интересуют объекты первых двух типов. Введем основные определения:

1. Целевой объект – это множество точек в линейном подпространстве, которые являются конечным пунктом траектории движения мобильной платформы для данной сцены.
2. Исходный объект – объект, связанный с исходным положением мобильной платформы в пространстве.
3. Объект-препятствие – набор особых точек на сцене, представленный на трехмерной карте в виде множества точек, которые не позволяют построить близкую к оптимальной траекторию для заданных исходных и целевых объектов.
4. Объекты (семантические ориентиры) рассматриваются как набор трехмерных точек в виртуальном пространстве с локально-плоской структурой и представлены в виде трехмерных Гауссовских функций. При инициализации метод SLAM не имеет информации о семантических маркерах на сцене. Для определения семантических ориентиров на сцене используются методы семантической

сегментации изображений и трехмерных облаков точек, а также эталонная база данных NYU Depth dataset NYUv2, которая содержит размеченный набор данных (RGB-D кадров) с сегментацией: свыше 200 тыс. изображений и 80 классов объектов.

Решение задачи SLAM не подразумевает получения полной информации об исследуемой среде на каждом этапе работы алгоритма локализации, поэтому в результате решения задачи планирования траектории не может быть определена оптимальная траектория на сцене. Вместо этого вводится понятие локально-оптимальной траектории, определяющей кратчайший путь от исходной точки до целевой точки для доступного фрагмента трехмерной сцены. После завершения процедуры обнаружения и классификации объектов на сцене, можно использовать информацию о принадлежности объектов к классам для разметки метрико-топологического графа на основе информации о семантических характеристиках пространства и построения локально-оптимальной траектории движения мобильной платформы в динамической среде с препятствиями.

Введем понятие метрико-топологического графа:

$$\text{MT-Graph} = \langle A, \text{dist} \rangle, \quad (5.18)$$

где A – матрица вида $A_{m \times n} = \{a_{ij}\}: a_{ij} = 0 \vee 1, \forall i, j: 0 \leq j < m, i: 0 \leq i < n; m, n \in N \setminus \{0\}$, dist – метрика на множестве $A^+ = \{a_{ij} | a_{ij} \in A, a_{ij} = 0\}$. Клетку метрико-топологического назовем «проходимой», если $a_{ij} = 0$, «непроходимой», если $a_{ij} = 1$. В качестве метрики dist будем использовать диагональную эвристическую функцию.

Для разметки метрико-топологического графа используется информация о габаритах объектов на сцене и информация о семантических свойствах объектов на анализируемой сцене. Использование семантической информации о свойствах предметной области позволяет генерировать более «разреженные» от препятствий матрицы A с одной стороны, а с другой стороны исключить из рассмотрения задачи планирования траектории ряд путей, которые связаны с тупиковыми или потенциально опасными для мобильной платформы направлениями движения,

либо наоборот включить некоторые пути, которые представляют потенциальный интерес в рамках решаемой поисковой задачи в системе спасения людей.

В работе для решения задачи планирования траектории был использован известный алгоритм HGA* (Hierarchical A* for MT-Graphs) [180], который был модифицирован введением в структуру метрико-топологического графа информации о семантических маркерах [23]. Давайте рассмотрим задачу нахождения локальной-оптимальной траектории мобильной платформы на двумерной сцене как задачу поиска пути на МТ-графе.

В результате работы HGA* алгоритма осуществляется поиск опорных клеток (частично-упорядоченное множества клеток) «в матрице А, при этом между двумя любыми опорными клетками путь на плоскости может быть получен стандартными методами, например, на основе алгоритма построения дискретной прямой Брезенхема [181]» [180].

Обобщенная схема работы HGA* алгоритма выглядит следующим образом:

«Шаг 1. Формирование частичного пути на графе PP_0 и множества частичных путей кандидатов $PC=\{PP_0\}$ для дальнейшего рассмотрения. Положим $i=1$.

Шаг 2. Выбор частичного пути PP_i из множества PC и двух опорных клеток a_{ij} и a_{kl} Построение нуль-траектории между данными клетками на основе алгоритма Брезенхема.

Шаг 3. Проверяется проходимость нуль-траектории: если нуль-траектория проходимая, то секция $\langle a_{ij}, a_{kl} \rangle$ размечается как проходимая. Если нуль-траектория непроходимая, то осуществляется поиск препятствия, расположенного ближе других к клетке a_{ij} . Далее формируются две проходимые секции, на основе которых строится два варианта расширения текущего частичного пути. Данные частичные пути заменяют текущий частичный путь PP_i , осуществляется выбор других клеток, $i=i+1$ и возврат на Шаг 2.

Шаг 4. Проверка: если все секции PP_i проходимые, то искомый путь найден, иначе осуществляется выбор других клеток, $i=i+1$ и возврат на Шаг 2» [180].

Приведем описание алгоритма построения множества путей в графе с использованием промежуточных вершин графа и алгоритма HGA*:

Шаг 1. Построить путь от исходного объекта (клетки) s до целевого объекта (клетки) g на плоскости с использованием алгоритма HGA*.

Шаг 2. Определить множество промежуточных вершин в метрико-топологическом графе MT-Gr.

Шаг 3. Найти путь от исходного объекта (клетки) s до целевого объекта (клетки) g с учетом промежуточных вершин на основе эвристической функции.

Шаг 4. Если построено заданное число путей, то завершить алгоритм, иначе вернуться к Шагу 2.

Далее представим результаты вычислительных экспериментов и компьютерного моделирования. Для экспериментов была использована камера Kinect 2.0 в качестве RGB-D датчика и коридор Челябинского государственного университета в качестве трехмерной сцены. Эксперимент может быть описан следующим образом: мобильная платформа движется (вперед и назад) и измеряет расстояние до объектов в пределах своей дальности (35 метров). Было захвачено 122 RGB-D ключевых кадра. Изображения были подвергнуты дискретизации, в результате были получены изображения с разрешением 640*480. Аналогично плотные облака точек уменьшаются до 0,1 размера воксельной сетки. Производительность предложенного алгоритма планирования траектории была сравнена с известными методами в подобных условиях в терминах точности и вычислительной сложности (Рис. 5.9 и Рис. 5.10).

Из рисунка 5.9 видно, что эвристический алгоритм планирования траектории выбирает гораздо более короткий путь (отмечен синим цветом), чем известный алгоритм планирования траектории (отмечен черным цветом) на основе динамического алгоритма A* [80] для сложных сцен с препятствиями. Для сцен с простой топологией объектов и препятствий (Рис. 5.10) предложенный алгоритм планирования траектории показывает подобные результаты с известным алгоритмом планирования траектории A* с точки зрения выбора лучшей локально-оптимальной траектории для заданной сцены.



Рисунок 5.9 – Диаграмма пути для предложенного и классического алгоритма планирования траектории (динамический алгоритм A^*) для сцен сложной топологией

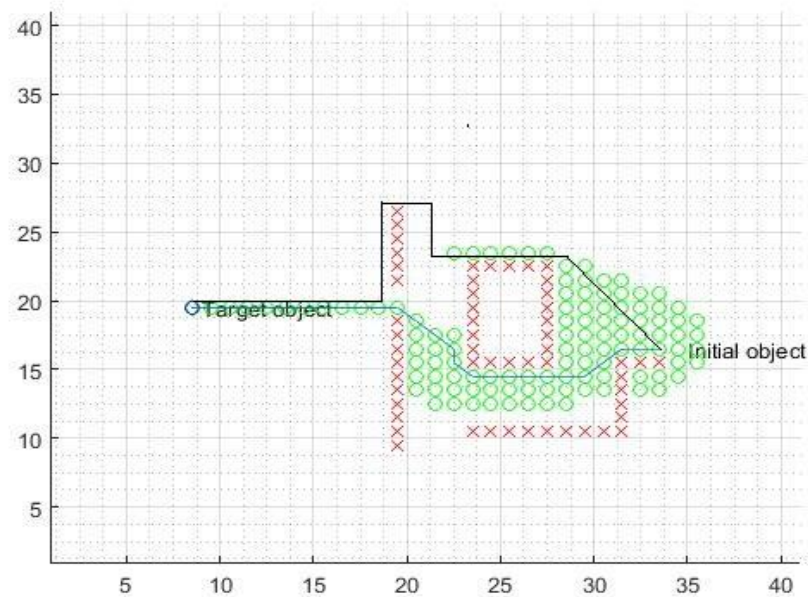


Рисунок 5.10 – Диаграмма пути для предложенного и классического алгоритма планирования траектории (динамический алгоритм А*) для сцен с простой топологией

Точность реконструкции траектории мобильной платформы и производительность предложенного алгоритма планирования траектории на основе метрико-топологического графа сравнивалась с известным алгоритмом, работающим в аналогичных условиях – динамический A^* + EKF SLAM, см. таблицу 5.4 и таблицу 5.5. Для проведения сравнительного анализа в работе использовались следующие базы данных NYU Depth Dataset V2 [179] и ASL Data Set [182], которые содержат семантическую маркировку пространства.

Таблица 5.4 – Точность реконструкции траектории мобильной платформы. Ошибка положения мобильной платформы для контролируемых условий сцены

Число навигационных ориентиров	EKF SLAM	S-EKF SLAM
Точность-ошибка положения, стандартное отклонение		
10	0.6	0.55
50	0.38	0.26
500	0.35	0.24
1500	0.32	0.21
5000	0.36	0.21
Время обработки, сек		
	241.5	173.6

В таблицах представлены оценки точности реконструкции траектории мобильной платформы по набору семантических маркеров на сцене в зависимости от числа навигационных ориентиров на исследуемой сцене в контролируемых (Табл. 5.4) и неконтролируемых условиях (Табл. 5.5).

Предложенный алгоритм планирования траектории дает наилучшую производительность и точность, чем известный метод планирования траектории на основе динамического A^* + EKF SLAM.

Таблица 5.5 – Точность реконструкции траектории мобильной платформы. Ошибка положения мобильной платформы для неконтролируемых условий сцены

Число навигационных ориентиров	EKF SLAM	S-EKF SLAM
Точность-ошибка положения, стандартное отклонение		
10	0.48	0.46
50	0.46	0.39
500	0.45	0.37
1500	0.43	0.35
5000	0.42	0.34
Время обработки, сек		
	386.3	221.5

5.5 Прототип системы на основе комбинированной методологии SLAM

В рамках диссертационного исследования разработан аппаратно-программный комплекс, состоящий из группы мобильных платформ. Для создания комплекса в предыдущих главах были решены следующие задачи:

1. Разработаны новые методы построения трёхмерной карты неизвестной окружающей среды с высокой точностью с использованием доступной мультисенсорной информации.
2. Разработаны алгоритмы навигации и согласованного функционирования мобильных платформ в каждый момент времени.
3. Исследовано современное состояние поставленной научно-технической проблемы, выполнен аналитический обзор научной литературы.
4. Обоснован выбор направления исследования, выбор методов и средств, а также предложены способы решения поставленных задач.
5. Выполнены патентные исследования в соответствии с ГОСТ Р 15.011-96.
6. Проведён анализ существующих алгоритмов калибровки и коррекции данных, получаемых от мультисенсорных датчиков.
7. Определены методы навигации мобильных платформ на сцене с использованием построенных трёхмерных карт окружающей среды.
8. Определён состав требований для лаборатории-полигона, предназначенной для проведения экспериментальных исследований.
9. Разработаны методы, алгоритмы для регистрации данных глубины относительно облака точек в трёхмерном пространстве на основе итеративного метода ближайших точек ICP в классе аффинных и ортогональных преобразований.
10. Разработан алгоритм для согласованного и оптимизированного планирования траекторий мобильных платформ с использованием построенной

трёхмерной карты окружающего пространства с учётом возможных статических и динамических препятствий.

11. Разработаны алгоритмы и программная реализация определения и отслеживания основных структурных элементов помещения (пол, потолок, стены).

12. Разработаны методы коллективного слежения за трехмерными объектами и вычисления их траекторий на искаженных динамических сценах с контролируемыми ошибками локализации на основе адаптивной фильтрации.

13. Создана интеграция разрабатываемой системы одновременной навигации и построения карты с мобильной платформой, используемой в макетном образце.

14. Разработаны алгоритмы управления и программная реализация пространственно-распределённой группой мобильных платформ.

15. Изготовлен экспериментальный образец (ЭО), состоящий из группы мобильных платформ в составе трёх колёсных сервисно-информационных платформ с мультисенсорными датчиками.

16. Разработана программа и методика экспериментальных исследований для группы мобильных платформ.

17. Проведены экспериментальные исследования ЭО группы мобильных платформ.

Разработанный ЭО решает широкий спектр прикладных задач с лучшим качеством (с точки зрения точности и производительности), чем известные существующие системы SLAM, работающие в аналогичных условиях.

В диссертационном исследовании проведены эксперименты с реальными данными и компьютерное моделирование на эталонных базах данных. Для проведения экспериментальных исследований была сформирована серия сценариев поведения мобильной платформы и создана тестовая база данных, содержащая изображения и трехмерные облака точек.

Разработанный программно-аппаратный комплекс выполняет следующие функции:

- осуществление навигации в неизвестном окружении;

- реконструкция окружающего пространства и распознавание трехмерных объектов по динамическим мультисенсорным данным на сценах;
- планирование траектории мобильной платформы в неизвестном динамическом пространстве;
- идентификация и локализация личности на трехмерной сцене;
- распознавание личности по лицу на основе предложенных методов реконструкции окружающего пространства.

Компоненты разработанной системы одновременной навигации и построения карты поддерживают работу на следующих сенсорах глубины:

- ASUS Xtion;
- Microsoft Kinect;
- времяпролетная камера MESA.

Исходный код предложенной системы одновременной навигации и построения карты переносим на следующие операционные системы: Linux и Android.

Исходный код разработанной системы для одновременной навигации и построения карты переносим на следующие архитектуры процессора: x86-64 и ARM.

Все компоненты программно-аппаратного комплекса обеспечивают режим работы в реальном времени с частотой не менее 30 кадров в секунду.

Входные данные для исследований – RGB/StereoRGB/RGBD изображения.

Требования к входным данным и объектам сцены:

- объекты на сцене имеют жесткую форму;
- при интеграции разрабатываемого программно-аппаратного комплекса с мобильной платформой не используется одометрическая информация;
- объекты в окружающем пространстве могут менять свое расположение;
- степень свободы движения сенсора на мобильной платформе 6D;
- план/карта внешнего окружения не заданы;

- во внешнем окружении присутствуют динамические объекты и субъекты;
- внутренние параметры используемых камер известны и не меняются в процессе работы системы SLAM.

Все входные данные для комбинированной методологии были получены с помощью датчиков: камера видимого диапазона Beward (Бевард, Россия) камера Kinect 2.0 (Microsoft, США) и времяпролетная камера MESA SR4000 (MESA Imaging, Швейцария). В работе проведен анализ существующих алгоритмов калибровки и коррекции данных, получаемых от различных камер.

Технические характеристики мультисенсорных датчиков:

- промышленная времяпролетная камера MESA SR4000, максимальная дальность сканирования: 5-10м, угол сканирования: $43^{\circ} \times 34^{\circ}$ (Ш/В), разрешение: 176x144 пикселей, угловое разрешение: $0,24^{\circ}$, частота обновления кадров: 33-100мс, размеры: 65x65x76мм;
- камера видимого диапазона Beward B272 с широким углом обзора (до 180 градусов), цифровой zoom (x) 400, высокое разрешение HD Ready1080i, режим день/ночь (DSS + механический ИК-фильтр), соответствие стандарту ONVIF, рабочая температура $-20...+50^{\circ}\text{C}$;
- камера видимого диапазона Beward B272 с узким углом обзора, с высоким разрешением HD Ready1080i, режим день/ночь (DSS + механический ИК-фильтр), соответствие стандарту ONVIF, рабочая температура $-20...+50^{\circ}\text{C}$.

5.6 Выводы по главе

1. В работе предложен метод семантической маркировки неизвестного окружающего пространства на основе Байесовского слияния метода классовой трехмерной сегментации 3DEF и комбинации метода семантической сегментации Grabcut и детектора объектов YOLOv3. Результаты семантической маркировки были применены при решении задачи регистрации трехмерных данных на основе решения вариационной задачи с метриками точка-плоскость и точка-точка.

2. Впервые был предложен метод решения задачи одновременной навигации и построения карты, основанный на комбинировании визуальных и семантических характеристик неизвестной окружающей среды, получены формулы для оценки вектора состояния при движении мобильной платформы по серии зашумленных измерений и семантических маркеров.

3. Предложенный метод решения проблемы S-EKF SLAM базируется на использовании мультисенсорной динамической информации, на пространственном согласовании последовательных кадров данных, на локализации семантических маркеров, на обнаружении «замыканий цикла», на согласовании полной последовательности кадров и глобальной оптимизации.

4. Предложенный метод может быть использован для приближенного решения систем уравнений и неравенств с заданными алгоритмическими функциями, удовлетворяющими условию Гельдера, что позволит создать точную систему одновременной локализации и отображения, работающую в режиме реального времени.

5. Был проведен сравнительный анализ методов в терминах точности и вычислительной сложности с известными открытыми системами SLAM HOGMan, ORB-SLAM, TORO-SLAM, RGBDSLAM, EKFMonoSLAM, которые используются в коммерческих системах Stanley (DARPA, Stanford Artificial Intelligence Lab, the Stanford Racing Team, Volkswagen ElectronicsResearch Laboratory), Robots Podcast: 3D SLAM (Department of Defense Research and Development Laboratory,

Massachusetts Institute of Technology (MIT) Lincoln Laboratory), Autonomous and Perceptive Systems research (University of Groningen) и Active 3-D Visual SLAM (General Electric Company). Полученные при выполнении работы результаты показывают, что предложенный комбинированный метод S-EKF SLAM показывает лучшие по точности результаты, чем известные подходы к решению проблемы SLAM на основе визуального SLAM в неконтролируемых условиях.

6. Полученные результаты компьютерного моделирования позволяют сделать вывод, что применение семантических характеристик окружающей среды позволяет улучшить точность и сходимость решения вариационной задачи ICP с метрикой точка-плоскость в контролируемых и неконтролируемых условиях наблюдения. Было установлено, что лучшие результаты по точности достигаются при комбинированном решении вариационной задачи, а основными факторами влияния на точность и сходимость методов регистрации в порядке значимости являются: совместное решение вариационной задачи на основе данных об особых точках и данных о глубине; использование семантической информации об характеристиках окружающей среды; решение проблемы зависимости точности результата метода регистрации от выбора начального приближения.

7. В работе рассмотрен эффективный эвристический алгоритм планирования траектории мобильной платформы в неизвестном динамическом пространстве, обладающий теоретически обоснованными оценками вычислительной сложности на основе метрико-топологического графа, содержащего семантическую информацию о свойствах окружающей среды. Применение семантической информации при поиске пути на двумерной сцене с использованием МТ-графа позволяет повысить производительность процедуры планирования траектории мобильной платформы. Алгоритм планирования траектории показывает лучшие результаты в терминах точности, чем известные алгоритмы планирования траектории для сцен со сложной структурой препятствий как в условиях наличия искажений окружающей среды, так и без них.

8. Основные результаты по данной главе опубликованы в работах [18, 23, 24, 29, 30, 32, 35, 41].

ГЛАВА 6 МЕТОДЫ ВОССТАНОВЛЕНИЯ И СЖАТИЯ ИЗОБРАЖЕНИЙ И СИГНАЛОВ

В данной главе рассмотрена задача предварительной обработки данных, связанная с восстановлением изображений от различных помех и сжатием изображений. Предложен быстрый рекурсивный алгоритм для вычисления дискретного косинусного преобразования (ДКП) для решения задачи сжатия изображений.

Рассмотрим современное состояние проблемы. Процесс получения цифровых изображений, как правило, сопровождается такими явлениями, как дефокусировка, размытие сигнала в результате относительного движения камеры и объекта, неравномерное освещение, шум различной природы. Среди типичных искажений изображений можно выделить искажения, связанные с оптической системой и электронными сенсорами, и искажения, связанные с состоянием среды, в которой находится регистрируемый объект. Поэтому задачи обнаружения, классификации объектов и слежения за объектами на статических и динамических сценах в большинстве случаев решаются в условиях помех и искажений. В данной главе рассмотрены методы компенсации этих явлений или восстановления изображений. Точность восстановления изображений существующими методами имеет определенные ограничения. Важную роль среди современных методов восстановления играют методы, основанные на применении вариационных функционалов [189]. Одним из широко применяемых методов численного решения вариационных задач является метод градиентного спуска и его вариации, такие как метод проекции градиента [190]. В последнее время развиваются методы, основанные на понятии расстояния Брегмана, пришедшего в теорию цифровых изображений из функционального анализа [191]. Одним из наиболее известных подходов к решению задачи восстановления зашумленного изображения является метод полной вариации [192]. В данной работе вариационная задача сформулирована, как задача условной минимизации. В работах [193, 194] дается

эквивалентная безусловная постановка задачи и доказываются существование и единственность решения. Большое количество работ посвящено распространению метода полной вариации на задачи восстановления изображения, искаженного не только аддитивным шумом, но и действием линейного оператора. Одним из подходов к разработке новых методов восстановления является обнаружение принципиальных недостатков существующих методов и их компенсация. Одним из шагов в этом направлении является применение к вариационным методам Фурье анализа [195]. Изучению влияния ядра линейного искажающего оператора на точность итерационных методов восстановления и на качество обнаружения объекта на восстановленном изображении с помощью оптимальной согласованной фильтрации посвящены работы [196, 197].

Для решения задачи восстановления изображений используется также метод, основанный на микросканировании матричных приемников [198-202]. Фильтрация и восстановление изображения выполняется путем решения системы линейных или нелинейных уравнений, которые получаются минимизацией функционала ошибки. Микросканирование выполнено с помощью контролируемых смещений камеры и с помощью контролируемых смещений источника света (для задачи устранения неравномерного освещения). Задача восстановления изображений от пространственно-неоднородных искажений может быть решена с помощью локально-адаптивных методов фильтрации [203-205]. Эти адаптивные линейные фильтры могут быть успешно реализованы в области ортогонального преобразования, которое определяется в каждом положении скользящего окна. Для реализации фильтров предложен быстрый метод синусоидального преобразования, использующий рекурсивные уравнения второго порядка [206, 207]. Полученные результаты восстановления изображений в исследовании сравнены с результатами известных методов по восстановлению и фильтрации изображений.

6.1 Рекурсивное вычисление дискретного косинусного преобразования для обработки сигналов с пониженным разрешением

В данной работе предлагается быстрый алгоритм для вычисления дискретного косинусного преобразования в локальном окне, скользящем по дискретному сигналу с шагом большим, чем единица [9]. Алгоритм основан на рекурсивном уравнении второго порядка между косинусными спектрами, которые вычисляются в равноудаленных локальных окнах. В работе [9] также предложен алгоритм быстрого обратного косинусного преобразования для скользящей обработки дискретных сигналов.

«Дискретное косинусное преобразование (ДКП) широко используется во многих приложениях, связанных с обработкой дискретных сигналов, таких как адаптивная фильтрация, видеообработка сигнала, выделение признаков и сжатие данных. Это объясняется тем, что ДКП действует на случайные марковские стационарные сигналы с коэффициентом корреляции между отсчетами близким к единице [208] подобно методу главных компонент (разложения Карунена-Лоэва). Во многих приложениях часто используется понятие локальной обработки сигнала с фильтрацией в области ортогонального преобразования [209].

Текущее ортогональное преобразование сигнала x_k может быть определено следующим образом [210]:

$$X_s^{kp} = \sum_{n=-N_1}^{N_2} x_{kp+n} \psi(n, s), \quad (6.1)$$

где $\{\psi(n, s)\}$ – базис ортогональных функций преобразования, k – целое число, p является целочисленным шагом скользящего окна и X_s^{kp} – ортогональные коэффициенты преобразования сигнала относительно дискретного времени kp » [9]. Общая схема локальной обработки сигнала в области ортогонального преобразования в скользящем окне представлена на рисунке 6.1.

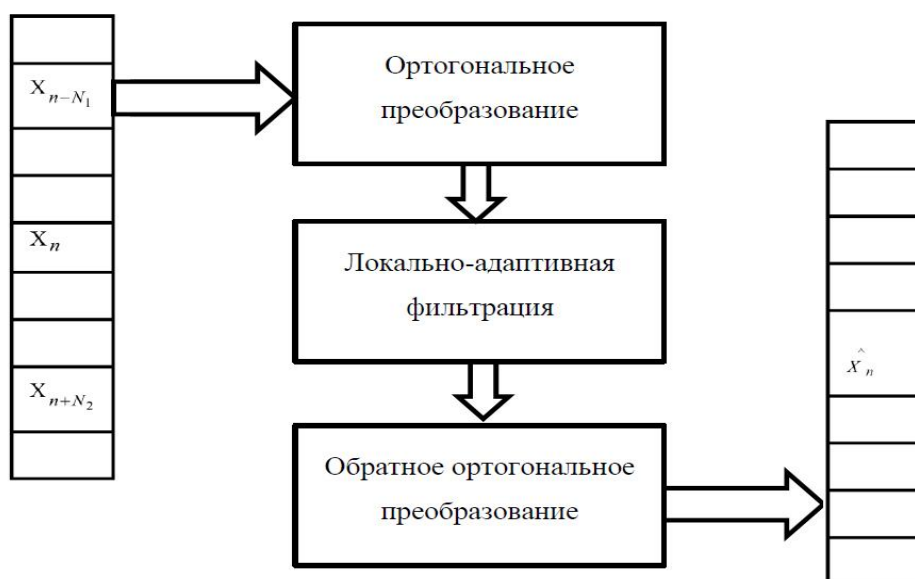


Рисунок 6.1 – Общая схема локальной обработки сигнала в области ортогонального преобразования в скользящем окне

Ядро ДКП порядка N можно выразить следующим образом

$$DCT_N = \left\{ k_s \cos \left(\pi \frac{s(n + 1/2)}{N} \right) \right\}, \quad (6.2)$$

где $n, s = 0, \dots, N-1, k_s = \begin{cases} 1/\sqrt{2}, & \text{если } s = 0; \\ 1 & \text{иначе.} \end{cases}$

«Если окно движется с шагом равным 1, можно использовать известные свойства смещения для дискретных синусоидальных преобразований для модификации коэффициентов преобразования. Свойства смещения первого порядка были описаны в работе [211]. Однако предложенный в этой работе алгоритм является не очень эффективным в связи с его высокой вычислительной сложностью. Поэтому были предложены быстрые алгоритмы для вычисления скользящего ДКП [212, 213] и рекурсивные уравнения второго порядка [214, 215] с шагом единица» [9].

6.2 Рекурсивный алгоритм для вычисления скользящего ДКП с произвольным шагом

«Текущее ДКП с шагом p может быть определено следующим образом

$$X_s^{kp} = \sum_{n=-N_1}^{N_2} x_{kp+n} \cos\left(\pi \frac{(n + N_1 + 1/2)s}{N}\right), \quad (6.3)$$

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. Опустим нормирующий множитель и особое масштабирование компоненты X_0^{kp} . Тогда 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» [9]

$$X_s^k = a_s X_s^{k-1} - X_s^{k-2} + f_s^k, \quad (6.4)$$

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 (6.4) – линейное неоднородное разностное уравнение второго порядка. Корни соответствующего характеристического уравнения [216, 217] равны $\lambda_s = \exp(ja_s)$ и $\gamma_s = \exp(-ja_s)$. После некоторых преобразований, общее решение разностного уравнения (6.4) может быть представлено в виде» [9]

$$X_s^k = \begin{cases} \frac{1}{k_1} \left(k X_s^{k_1} + (k - k_1) \left(\sum_{m=1}^{k_1-1} m f_s^{k_0+m+1} - X_s^{k_0} \right) \right) + \sum_{m=1}^{k-k_1} m f_s^{k-m+1}, & s = 0 \\ \frac{X_s^{k_1} \sin(a_s k) + \left(\sum_{m=1}^{k_1-1} f_s^{k_0+m+1} \sin(a_s m) - X_s^{k_0} \right)}{\sin(a_s) \sin(a_s k_1)} * \sin(a_s (k - k_1)) + \sum_{m=1}^{k-k_1} f_s^{k-m+1} \frac{\sin(a_s m)}{\sin(a_s)}, & s = 1, \dots, N-1 \end{cases}, \quad (6.5)$$

«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 $k > k_1$.

Рекурсивное соотношение между тремя спектрами: $X_s^{(k-2)p}$, $X_s^{(k-1)p}$ и X_s^{kp} может быть получено из уравнения (6.5) следующим образом»[9]

$$X_s^{kp} = \begin{cases} 2X_s^{(k-1)p} - X_s^{(k-2)p} + \sum_{m=1}^{p-1} (f_s^{(k-2)p+m+1} + f_s^{kp-m+1})m + f_s^{(k-1)p+1}p, s = 0 \\ 2X_s^{(k-1)p} \cos(a_s p) - X_s^{(k-2)p} + \sum_{m=1}^{p-1} (f_s^{(k-2)p+m+1} + f_s^{kp-m+1}) \frac{\sin(a_s m)}{\sin(a_s)} + f_s^{(k-1)p+1} \frac{\sin(a_s p)}{\sin(a_s)}, s = 1, \dots, N-1 \end{cases} \quad (6.6)$$

«Из соотношения (6.6) видно, что при вычислении скользящего ДКП используются значения входного сигнала x_k , а также коэффициенты ДКП, вычисленные в двух предыдущих положениях скользящего окна. Если $p=1$, то рекурсивное уравнение упрощается до уравнения (6.4).

Количество арифметических операций, требуемых для вычисления ДКП при $p>1$ в данном положении окна, оценивается следующим образом: если N – нечетное, то вычислительная сложность составляет $N(p+1)$ операций умножения и $N(p+1)+5(p-1)$ операций сложения; если N – четное, количество операций умножения уменьшается до $N(p+3)/2$. Таким образом, общее количество необходимых арифметических операций может быть оценено следующим образом: $2N(p+1)+5(p-1)$ и $\frac{N}{2}(3p+5) + 5(p-1)$, когда N является нечетным и четным соответственно.

Рекурсивный алгоритм ДКП [212, 213] требует приблизительно $4Np$ арифметических операций для вычисления спектра ДКП в равноотстоящих положениях окна kp . Можно отметить, что предложенный алгоритм при $p>1$ является более эффективным, чем рекурсивный алгоритм ДКП. Таблица 6.1 показывает количественные результаты тестирования вычислительной сложности для предложенного ДКП, быстрого ДКП [218, 219] и рекурсивного алгоритма ДКП [212, 213] для $N=256$. Из таблицы 6.1 видно, что предлагаемый алгоритм ДКП дает

лучшие результаты, когда шаг скользящего окна p меньше, чем граничное значение, равное 8» [9].

Таблица 6.1 – Сравнение вычислительной сложности алгоритмов для вычисления скользящего ДКП (плавающие операции сложения и умножения), $N=256$

p -шаг скользящего окна	Алгоритмы		
	Быстрое дискретное косинусное преобразование	Рекурсивное дискретное косинусное преобразование	Предлагаемое дискретное косинусное преобразование
1	3708	1024	1024
2	3708	2048	1413
3	3708	3072	1802
4	3708	4096	2191
5	3708	5120	2580
6	3708	6144	2969
7	3708	7168	3358
8	3708	8192	3747
9	3708	9216	4136

Граничные значения для шага скользящего окна в зависимости от длины окна N приводятся в таблице 6.2.

Таблица 6.2 – Граничные значения шага p для предлагаемого алгоритма в зависимости от длины скользящего окна N

N	16	32	64	128	256	512	1024
P	3	4	6	7	8	9	10

«Длина скользящего окна для предлагаемого алгоритма может быть произвольным целочисленным значением, определяемым характеристиками сигнала, требующего обработки» [9].

6.3 Точность вариационных методов восстановления сигналов, искаженных линейным оператором

Среди типичных искажений изображений можно выделить искажения, связанные с оптической системой и электронными сенсорами, а также искажения, связанные с состоянием среды, в которой находится регистрируемый объект. Применение вариационных методов для восстановления изображений исследовано в работах Кобера В.И., Маковецкого А.Ю., Диас-Рамиреса В.Х, Агилара-Гонсалеса П.М. [190-192]. В данном параграфе предлагается метод восстановления изображений, который основывается на способе решения вариационной задачи, предложенном в перечисленных работах, и описывается влияние ядра линейного искажающего оператора на точность итерационных методов восстановления и на качество обнаружения объекта на восстановленном изображении с помощью оптимальной согласованной фильтрации [12]. Процесс получения цифровых изображений в большинстве случаев сопровождается такими явлениями, как дефокусировка, смаз в результате движения, зашумление различной природы.

При восстановлении изображений часто возникает необходимость решить следующую обратную задачу

$$A \tilde{u} = u_0, \quad (6.7)$$

где A – известный линейный оператор, u_0 – наблюдаемое изображение, $\tilde{u}$ и $u_0 \in U(\Omega)$, $U(\Omega)$ – банахово пространство на ограниченном множестве Ω . Функция $\tilde{u}$ называется точным решением задачи (6.7). В большинстве случаев обратный оператор A^{-1} не существует, а наблюдаемое изображение u_0 может быть дополнительно искажено аддитивным шумом n :

$$v_0 = u_0 + n. \quad (6.8)$$

Тогда задача (6.7) принимает вид

$$Au = v_0. \quad (6.9)$$

Типичным подходом к решению этих задач является использование вариационного функционала

$$J(u) = H(Au) + \lambda R(u), \quad (6.10)$$

где $H(Au)$ называется точностным членом, а $R(u)$ – регуляризирующим членом, $\lambda > 0$ – параметр регуляризации. Часто используемая следующая форма точностного члена

$$H(Au) = \| Au - u_0 \|_{L_2}^2. \quad (6.11)$$

Для получения приближенного решения задачи (6.9) рассматривается следующая вариационная задача

$$u = \arg \min \| Au - u_0 \|_{L_2}^2 + \lambda R(u). \quad (6.12)$$

Одним из наиболее широко используемых регуляризирующих членов является полная вариация [195, 196]

$$R(u) = \int_{u \in BV(\Omega)} |\nabla u|. \quad (6.13)$$

Задаче (6.7) соответствует вариационная задача (6.12) без регуляризирующего члена

$$u = \arg \min \| Au - u_0 \|_{L_2}^2. \quad (6.14)$$

Для решения задачи (6.14) часто используется метод градиентного спуска:

$$u_{k+1} = u_k - \alpha_{k+1} \cdot H'(u_k), \quad (6.15)$$

где $H'(u)$ – градиент функционала $H(Au)$, α_k – параметр градиентного метода. В функциональном пространстве L_2 градиент задается формулой

$$H'(u_k) = 2A^*(Au - u_0), \quad (6.16)$$

где A^* – сопряженный оператор.

Параметр α_k выбирается из условия

$$H(u_{k+1}) < H(u_k). \quad (6.17)$$

В работе [12] показано, что функция u_{k+1} , полученная на $k + 1$ итерации градиентного метода, может быть выражена следующим образом

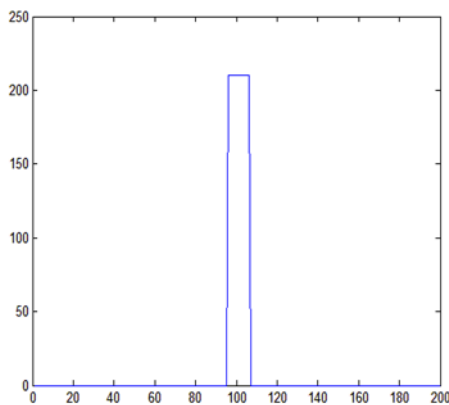
$$U_{k+1} = \left(\begin{array}{c} (1 - 2\alpha A^2)^{k+1} A + \\ + ((1 - 2\alpha A^2)^k + \dots + (1 - 2\alpha A^2)^1 + 1) 2\alpha A^2 \end{array} \right) (\tilde{u}). \quad (6.18)$$

Фурье-образы функций u_{k+1} и $\tilde{u}$ связаны следующим уравнением

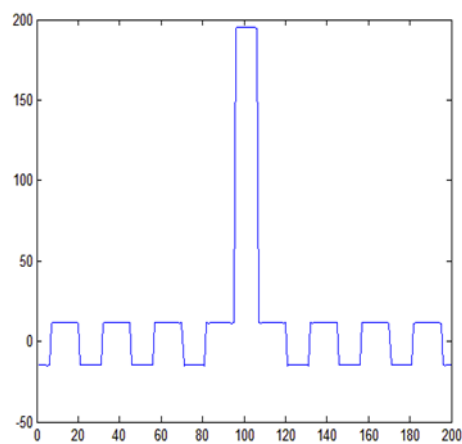
$$\mathcal{F}(u_{k+1}) = (1 - (1 - \text{sinc}(2\Delta f))(1 - 2\alpha \text{sinc}^2(2\Delta f))^{k+1}) \cdot \mathcal{F}(\tilde{u}). \quad (6.19)$$

Таким образом, в точках $f = \frac{\pi n}{2\Delta}$ функция $\mathcal{F}(u_k) = 0$ для любого шага k градиентного метода. Следовательно, в этих точках Фурье-образ восстановленной функции имеет нулевые значения независимо от значений Фурье-образа исходной функции в этих точках.

На рисунке 6.2, а) показана исходная дискретная функция $v(t)$.



а)



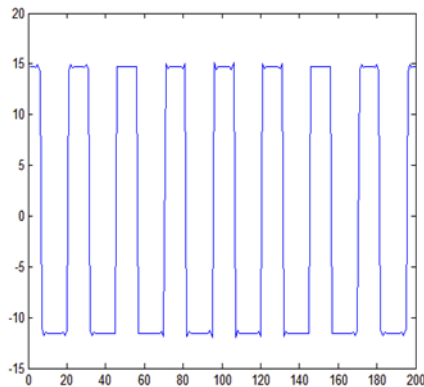
б)

Рисунок 6.2 – Дискретная функция $v(t)$: а) тестовая дискретная функция, б) восстановленная функция

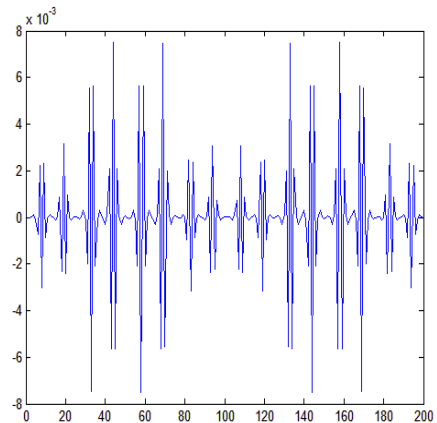
Ядром линейного оператора $\mathcal{A}: X \rightarrow Y$, X и Y – банаховы пространства, называется множество $\text{Ker } \mathcal{A} = \{x \in X \mid \mathcal{A}x = 0\}$. Предположим, что функция $\tilde{u}$ является решением задачи (6.14). Тогда функция $\tilde{u} + u_{\text{ker}}$ также является решением задачи (6.16), где $u_{\text{ker}} \in \text{Ker } A$. Существуют ли решения задачи (6.14), полученные методом градиентного спуска, которые отличаются от функций вида $\tilde{u} + u_{\text{ker}}$? Ответом на этот вопрос является следующая доказанная теорема [12].

Теорема. Метод градиентного спуска асимптотически сходится к функции вида $\tilde{u} + u_{\text{ker}}$.

На рисунке 6.3, а) показана разность между исходной и восстановленной функциями, на рисунке 6.3, б) показана функция $A(v - u_k)$.



а)



б)

Рисунок 6.3 – Графическая интерпретация теоремы: а) Разность между исходной и восстановленной функциями, б) функция $A(v - u_k)$

Таким образом, при решении задачи (6.14) градиентный метод сходится к функциям вида $v + u_{kl}$. Если ядро нетривиальное, т.е. содержит ненулевые элементы, то, как показывает рассмотренный пример, u_{kl} может быть одним из ненулевых элементов ядра.

6.4 Компьютерное моделирование

Введем сокращения для аддитивной модели формирования изображений и мультипликативной модели (ММ) формирования изображения. Введем следующие обозначения: АП – аддитивная помеха; НАП – неоднородная аддитивная помеха; МП – мультипликативная помеха, ММП – неоднородная мультипликативная помеха; АШ – аддитивный шум. Кобер В.И. предложил метод решения задачи восстановления изображений, основанный на микросканировании матричных приемников [198-202]. В данном параграфе представлены результаты компьютерного моделирования непараметрического метода фильтрации и восстановления изображений [39] (Кобер В.И., Вохминцев А.В.), проведен сравнительный анализ с известными методами:

1. На основе количественного критерия – среднеквадратичной ошибки (СКО).
2. На основе качественного критерия, который определяет визуальное качество восстановления сигнала и его степень устойчивости к разного рода шумам, например, позиционным ошибкам матричного сенсора (приемника).

СКО может быть определена по следующей формуле

$$RMSE = \sqrt{\frac{\sum_{k,l}^M (f(k,l) - \tilde{f}(k,l))^2}{M^2}}, \quad (6.20)$$

где $f(k, l)$ – исходный сигнал; $\tilde{f}(k, l)$ – восстановленный сигнал. Для проведения компьютерного моделирования используются изображения размером 256×256 . При компьютерном моделировании в тестах диапазон сигнала изменяется следующим образом: $[0, 255]$. Для получения корректных результатов была проведена серия тестов, а именно каждый типовой эксперимент был повторен в 50-ти различных экспериментах с различными значениями параметров входного шума. Для решения системы алгебраических уравнений при проведении компьютерного моделирования применяются итерационные методы Крыловского, например, алгоритм сопряженного градиента (Conjugate gradient method, CGM). В

серии тестов было установлено условие для остановки данного алгоритма: пороговое значение для уровня сигнала равно 10^{-10} . Давайте выберем визуальный критерий EDIF следующим образом

$$EDIF = c_1 (f(k, l) - \tilde{f}(k, l)) + c_2, \quad (6.21)$$

где c_1 и c_2 – константы нормализации визуального критерия. Положим для серии тестов следующие значения: $c_1 = 4$, $c_2 = 128$ для АМ и $c_1 = 1$, $c_2 = 128$ для (ММ) соответственно.

В серии тестов сигналы не являются однородными в пространстве, поэтому нельзя использовать метод линейной оптимальной фильтрации. Поэтому в качестве эталонного метода для сравнительного анализа в диссертационной работе используется Винеровская фильтрация [220]. При этом для ММ формирования изображения было осуществлено поэлементное логарифмирование исходного изображения, при этом аддитивный шум игнорируется. Таким образом, ММ соответствует АМ в данном случае. При моделировании Винеровская фильтрация была применена к восстановленному сигналу. Далее была выполнена операция поэлементного потенцирования к полученному результату. Результаты компьютерного моделирования представлены на рисунке 6.4: а) исходный сигнал (изображение), б) НАП, в) исходное изображение, искаженное помехами в виде НАП и белого шума (значение отклонения равно 2).

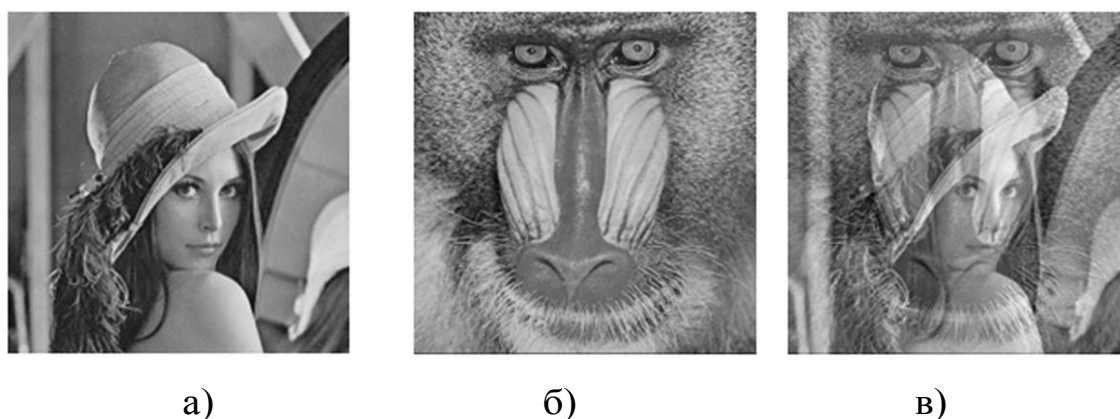


Рисунок 6.4 – Применение методов восстановления сигнала от искажений:

а) исходное изображение, б) НАП, в) исходное изображение, искаженное помехами в виде НАП и белого шума

В данном тесте «среднее значение» равняется 130 и СКО изображения-помехи равняется 40. На рисунке 6.5, а) представлен результат восстановления сигнала с помощью предложенного вариационного метода восстановления сигнала для АП. На рисунке 6.5, б) показан визуальный критерий EDIF для данного теста. Рассмотрим пример восстановления в СКО: для последовательности состоящей из трех исходных изображений «Am3»; для последовательности состоящей из пяти исходных изображений «Am5»; для Винеровской фильтрации, заданной в виде функции стандартного отклонения АШ.



Рисунок 6.5 – Применение вариационных методов восстановления сигналов для АП: а) результат восстановления сигнала с помощью предложенного метода восстановления, б) визуальный критерий EDIF для данного теста

Из рисунка 6.6 видно, что предложенные в работе вариационные методы показывают значительно лучшие результаты восстановления изображений в терминах среднеквадратичной ошибки в сравнении с известными методами на основе Винеровского фильтра. Отметим, что АП не являются пространственно-однородными, поэтому использование Винеровской фильтрации в данном случае некорректно. Проведем компьютерное моделирование для МНП на примере неравномерного освещения сцены. Данный вид помехи часто встречается при решении задачи реконструкции трехмерной реконструкции сцены и может существенно повлиять на точность предложенной в главе 2 комбинированной методологии SLAM.

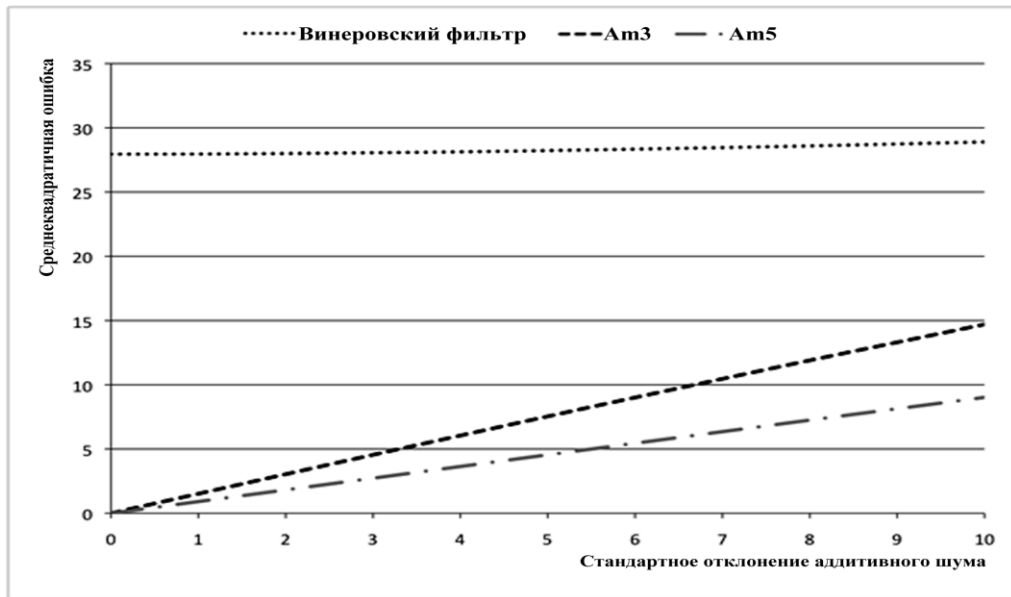


Рисунок 6.6 – Сравнение методов восстановления изображений в терминах СКО для АП

Результаты компьютерного моделирования представлены на рисунке 6.7.



а)

б)

в)

Рисунок 6.7 – Восстановление изображений от МНП на примере неравномерного освещения сцены для диапазона $[0.1, 1]$: а) вид функции освещенности для диапазона $[0.1, 1]$, б) исходное изображение сцены, в) изображение, которое было искажено МНП и белым шумом (значение отклонения равно 1)

На рисунке 6.8 показаны результаты восстановления в терминах СКО: для последовательности из трех исходных изображений «Mm3»; для

последовательности из пяти наблюдаемых изображений «Mm5»; для Винеровской фильтрации, заданной в виде функции среднеквадратической ошибки АШ в условиях неравномерного освещения.



а)



б)

Рисунок 6.8 – Применение вариационных методов восстановления сигналов для МП: а) результат восстановления сигнала, б) визуальный критерий для данного теста

Из рисунка 6.9 видно, что предложенные в диссертационной работе вариационные методы показывают значительно лучшие результаты восстановления изображений в терминах СКО в сравнении с известными методами на основе Винеровского фильтра с известными параметрами.

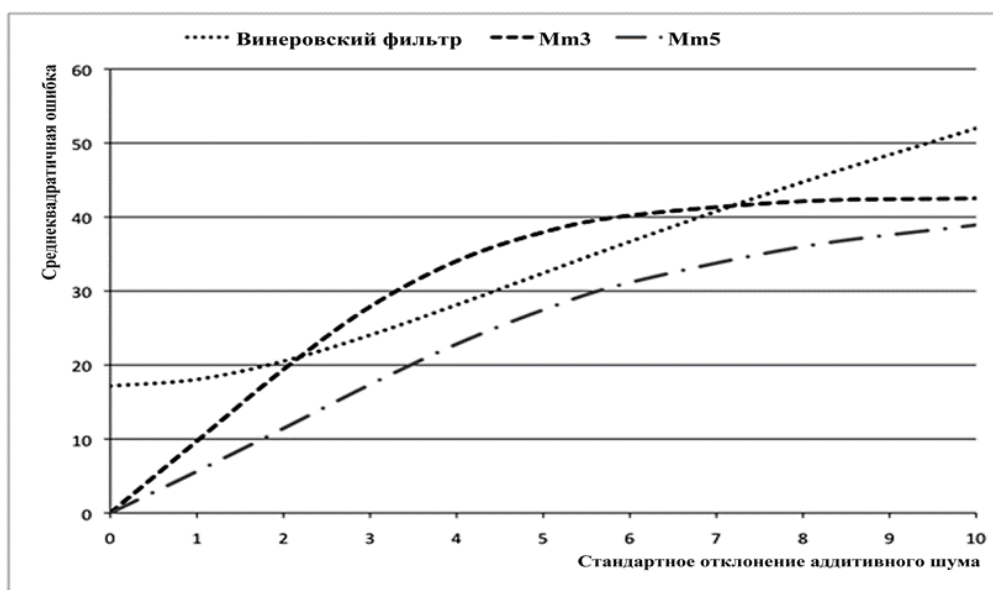


Рисунок 6.9 – Сравнение методов восстановления изображений в терминах СКО для МП

6.5 Модель базы данных для хранения информации о трехмерной сцене

Для автоматизации процесса обработки информации о трехмерной сцене и навигационных ориентирах и приведении данной информации к определенному формату, в работе создано программное обеспечение для конвертации результатов компьютерного моделирования в форматах las, inc, dev, csv к соответствующей реляционной структуре и для их загрузки в БД [28]. На рисунке 6.10 показана блок-схема алгоритма для обработки inc файлов с результатами компьютерного моделирования трехмерной сцены и для приведения к требуемой структуре БД.

В работе спроектирована реляционная БД, предназначенная для хранения и обработки информации о цвете, глубине и семантических характеристиках анализируемой трехмерной сцены. Дополнительно в работе был проведен анализ для устранения неполной и противоречивой информации о траектории мобильной платформы и координатах навигационных ориентиров на трехмерной сцене.

Ключевой таблицей в реляционной БД является таблица Well, которая содержит информацию об навигационных и семантических ориентирах, включая название, принадлежность к классу предметной области, связи с другими ориентирами в иерархической структуре объектов на сцене (3DEF – дерево), размеры ориентира, координаты на трехмерной сцене. Также для каждого навигационного ориентира на трехмерной сцене в БД имеется множество наборов данных, каждый из которых связан результатами компьютерного моделирования или вычислительного эксперимента (таблица WorkInfo). Информация в таблице представлена в хронологическом порядке.

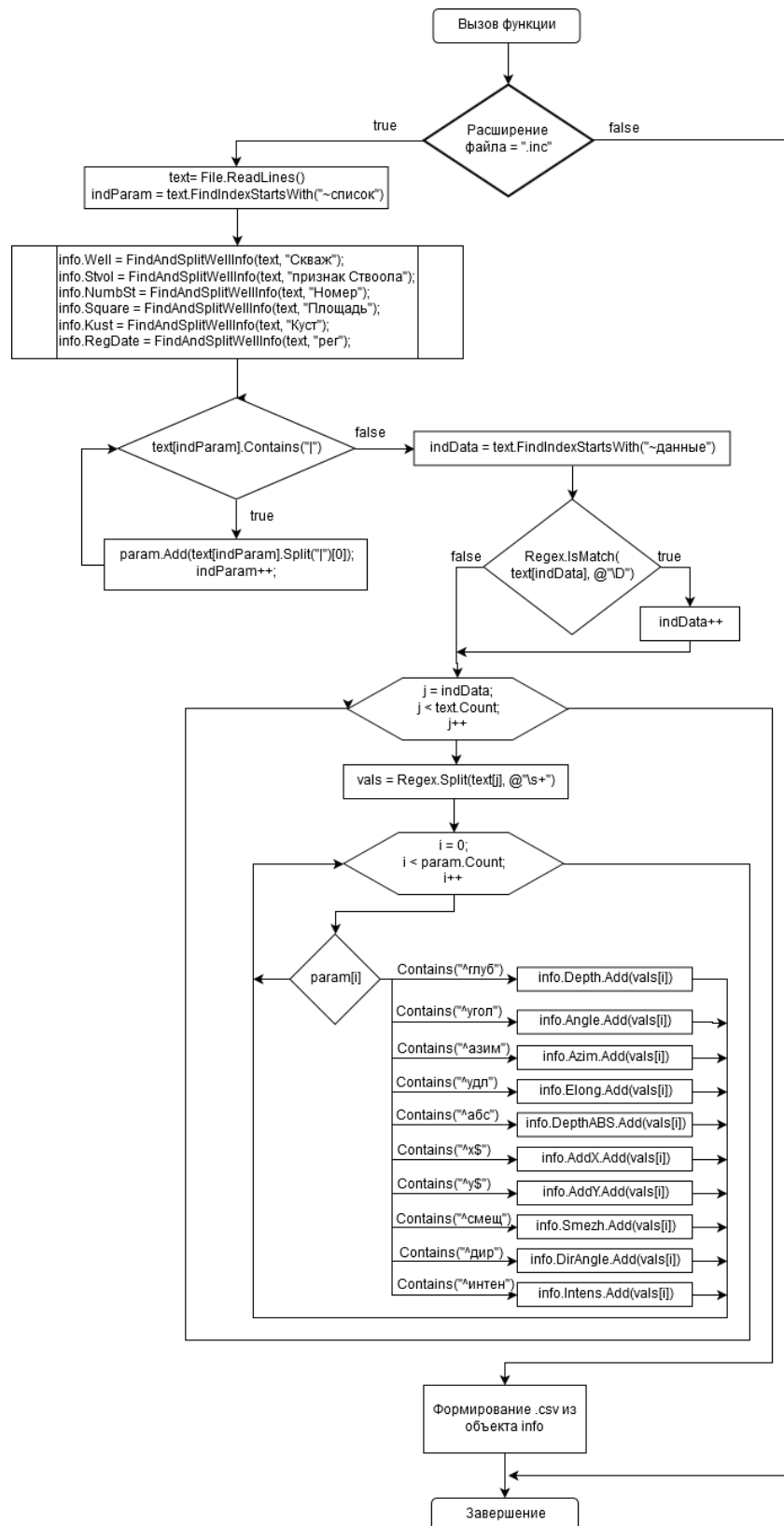


Рисунок 6.10 – Блок-схема алгоритма обработки файлов с результатами компьютерного моделирования трехмерной сцены

6.6 Выводы по главе

1. «В данной работе предложен 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. В работе приведены граничные значения для шага скользящего окна, при которых предлагаемый алгоритм является более эффективным, чем быстрые алгоритмы ДКП» [9].

2. Изучено влияние ядра линейного искажающего оператора на точность итерационных методов восстановления и на качество обнаружения объекта на восстановленном изображении с помощью оптимальной согласованной фильтрации. Представлен непараметрический метод фильтрации и восстановления изображений, искаженных аддитивным фоновым шумом, импульсными помехами и мультипликативными интерференциями, из последовательности пяти наблюдаемых изображений, полученных микросканированием камеры.

3. В точках $f=\pi n/2\Delta$ функция $\mathcal{F}(u_k) = 0$ для любого шага k градиентного метода. Следовательно, в этих точках Фурье-образ восстановленной функции имеет нулевые значения независимо от значений Фурье-образа исходной функции в этих точках.

4. Основные результаты по данной главе опубликованы в работах [9, 12, 28, 39].

ГЛАВА 7. РЕШЕНИЕ ПРОБЛЕМЫ «ЗАМЫКАНИЯ ЦИКЛА» И МЕТОДЫ ПОИСКА ИЗОБРАЖЕНИЯ ПО ВИЗУАЛЬНОМУ ПОДОБИЮ

В данной главе предложен 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.

Рассмотрим современное состояние проблемы. В результате решения проблемы SLAM определяется траектория движения мобильной платформы на сцене. Для повышения сходимости предложенного метода SLAM необходимо исключить из полученного решения участки с «замыканием цикла». Современные решения данной проблемы основаны на использовании методов быстрого сопоставления данных, которые пришли в робототехнику из области обработки текстов на естественном языке (Natural Language Processing) [61, 62]. Наиболее известные методы решения этой проблемы основаны на использовании корзины слов (Bag of Words) [54, 61, 67, 110]. Траектория движения мобильной платформы представляет некоторый связанный граф положений. При обнаружении на траектории движения «замыкания цикла», так называемой петли, создается новое ограничение в графе положений, которое определяет новое отношение между последовательностью RGB-D кадров для данной сцены. Для решения задачи одновременной оптимизации графа положений и карты сцены используется метод слепой подстройки [76] и его различные модификации [77-79]. Главным недостатком подхода корзина слов является низкая производительность, для преодоления которого был предложен метод поиска изображений по визуальному подобию с использованием «гистограммных слов» в древовидной структуре корзины и инвертированного индекса. В работе предложенный метод поиска по визуальному подобию использован для решения проблемы «замыкания цикла».

7.1 Поиск изображений по визуальному подобию с применением инвертированных индексов цветовых гистограмм

Рассмотрим современное состояние проблемы. «При решении задачи поиска изображений по визуальному подобию [221] в качестве признаков выбирают такие характеристики изображения, которые обеспечивают наилучшее обобщение графической информации в контексте выбранного понятия сходства: цветовые характеристики [222], особые точки [95-97], характеристики текстур [223], а также комбинации этих признаков [224]. В данной главе будет представлен метод поиска изображений по визуальному подобию для динамически пополняемых коллекций изображений с использованием метода индексации гистограмм изображений» [16].

«Пусть задана коллекция различных изображений $I = \{a_i\}_{i=1}^M$, для которой необходимо установить соответствие на основе оценки визуального сходства. В работе каждое изображение $a \in I$ характеризуется вектором признаков $\Phi(a) = \langle \varphi_1^a, \dots, \varphi_N^a \rangle$, где N – мощность пространства признаков $\mathbb{R}^N$. Два изображения $a \in I$ и $b \in I$ будем называть визуально схожими, если для выбранной функции оценки сходства $\rho(a, b)$ выполнено $\rho(a, b) = 1 - f(\Phi(a), \Phi(b)) \geq f_{min}$, где $f(\dots)$ – некоторая метрика в $\mathbb{R}^N$, а f_{min} – пороговое значение сходства.

Представляют интерес признаки, которые позволяют выявлять нечёткие дубликаты эталонных изображений [225]. Исходное изображение есть множество точек $P(a) = \{(x, y)\}$, для каждой точки задано значение цветности $C(x, y)$ в выбранном цветовом пространстве $\Omega: C(x, y): P(a) \rightarrow \Omega$. Для того чтобы идентифицировать незначительные изменения, исходное изображение в цветовом пространстве RGB приводится к цветовому пространству меньшей размерности $\sigma = |\Omega|$. Тогда гистограмма H_σ изображения $a \in I$ – есть отображение вида:» [16]

$$H_\sigma(a): P(a) \rightarrow \mathbb{R}^\sigma, \quad (7.1)$$

где $H_\sigma(a) = \langle v_1^a, \dots, v_\sigma^a \rangle$, $v_i^a = \frac{|\{(x, y) \in P(a) | C(x, y) = c_i\}|}{|P(a)|}$, c_i – цветовая компонента в цветовом пространстве размерности σ .

«Известно, что гистограммы различных по содержанию изображений могут быть близки. Для того чтобы учесть распределение цветов в разных областях изображения, будем рассматривать не гистограммы изображения в целом, а гистограммы отдельных его сегментов. Каждое изображение a сегментируется на области $S(a) = \{s_j(a)\}_{j=1}^K$ с возможным перекрытием областей.

Каждый сегмент $s_j(a)$ является квадратной областью исходного изображения, а $K = r \times r$, где r – параметр разбиения. Сегменты нумеруются последовательно слева направо, сверху вниз. Для каждого сегмента $s_j(a)$ строится гистограмма $H_\sigma(s_j(a))$, сравнение изображений выполняется посегментно» [16]

$$\varphi_i^a = v_{i-J\sigma}^{s_{J+1}(a)}, \quad (7.2)$$

«где $i = \overline{1, (r \cdot \sigma)}$, $J = \left\lfloor \frac{i}{r} \right\rfloor$, $[\cdot]$ – операция взятия целой части. Для оценки сходства изображений могут применяться метрики Минковского, Бхаттачария, косинусное и евклидово расстояния и др. Вычислительная сложность предложенного подхода методом прямого сравнения изображений по принципу «каждое с каждым» является квадратичной $O\left(\frac{I^2}{2}\right)$. Операция оценки сходства двух изображений реализуется на упорядоченных массивах значений признаков за один проход по этим массивам, её вычислительная сложность равна $O(r \cdot \sigma)$ » [16].

Недостатком предложенного метода является его вычислительная сложность, поэтому данный метод невозможно использовать при решении проблемы SLAM для крупномасштабных сцен, где требуется обрабатывать большие потоки данных.

7.2 Инвертированный индекс гистограмм для эффективного поиска изображений по визуальному подобию

«Для решения поставленной задачи быстрого поиска в области цифровой обработке данных применяют специальные структуры данных – инвертированные индексы (ИИ) [16], в которых производится предварительная обработка данных в соответствии с некоторым набором критериев [226]. Для построения ИИ изображения a по гистограммам его сегментов рассмотрим вектор $\Phi(a)$ и сопоставим гистограммные слова компонентам этого вектора. Гистограммные слова содержательно характеризуют исходные вещественные числовые значения φ_i^a . Введём конечное покрытие отрезка $[0;1]$ подотрезками

$$\Delta = \{[x_i; y_i] \mid 0 \leq x_i < y_i \leq 1, \bigcup_{i=1}^{|\Delta|} [x_i; y_i] = [0; 1]\}. \quad (7.3)$$

Разделим отрезок $[0;1]$ на G отрезков длины α , с перекрытием $\beta < \frac{\alpha}{2}$.

Поставим в соответствие каждой компоненте $\varphi_j^a, j = \overline{1, N}$ вектора $\Phi(a)$ множество упорядоченных пар $\xi_j^a = \{(j, i) \mid \varphi_j^a \in [x_i; y_i], i = \overline{1, |\Delta|}\}$: »[16]

$$E: \mathbb{R} \rightarrow \Lambda, \quad (7.4)$$

где $\Lambda = \{j \in \mathbb{N} \mid j \leq N\} \times \{i \in \mathbb{N} \mid i \leq |\Delta|\}$, т.е. $\xi_j^a \subseteq \Lambda$.

«Назовем упорядоченные пары вида $\langle j, i \rangle$ «гистограммными словами», где i представляет собой натуральные индексы отрезков $[x_i; y_i] \in \Delta$ таких, что значение φ_j^a попадает в эти отрезки. В определении соответствия E учтено, что значение φ_j^a может попадать в несколько перекрывающихся отрезков из покрытия Δ »[16].

Далее рассмотрим отображение $\Sigma: \mathbb{R}^N \rightarrow \Lambda$. По построению множеств ξ_j^a имеем, что $\xi_j^a \cap \xi_{j'}^a = \emptyset, j = j'$ и можно положить $\Sigma \Phi(a) = \bigcup_{j=1}^N \xi_j^a$. Данное отображение ставит в соответствие вектору признаков $\Phi(a)$ множество гистограммных слов.

«Введём отношение линейного порядка Π_1 на множестве Λ : пара $\langle j, i \rangle$ предшествует паре $\langle j', i' \rangle$ тогда и только тогда, когда:

1. Выполнено неравенство $j < j'$.
2. Выполнено равенство $j = j'$, но при этом $i < i'$.

На основе гистограммных слов строится ИИ изображения a . Для этой цели по предварительно построенному вектору $\Phi(a)$ выполняется построение прямого индекса изображения: $DI(a) = \{\langle j, i, \varphi_j^a \rangle | \langle j, i \rangle \in \sum \Phi(a)\}$. Затем данный индекс упорядочивается в соответствии с отношением порядка Π_1 по гистограммным словам $\langle j, i \rangle$.

ИИ представляет собой соответствие, где гистограммному слову $\langle j, i \rangle$ поставлено в соответствие линейно-упорядоченное множество: $II(\langle j, i \rangle) = \{\langle id^a, \varphi_j^a \rangle | a \in I\}$, где id^a – натуральный числовой идентификатор изображения a . ИИ строится с использованием слияния линейно-упорядоченных прямых индексов $DI(a)$ »[16].

На этапе поиска визуально похожих изображений выполняется следующая процедура:

Шаг 1. Построение вектора признаков $\Phi(c)$ по заданному эталонному изображению c .

Шаг 2. Формируется прямой индекс изображения $DI(c)$, гистограммные слова которого ($\langle j, i \rangle \in \sum \Phi(c)$) выступают в роли запроса к ИИ.

Шаг 3. Полученные для каждого гистограммного слова множества $II(\langle j, i \rangle)$ и $\langle j, i \rangle \in \sum \Phi(c)$ преобразуются во множества троек: $S(\langle j, i \rangle) = \{\langle id^a, j, \varphi_j^a \rangle | a \in I\}$, $\langle j, i \rangle \in \sum \Phi(c)$, добавлением в $\langle id^a, \varphi_j^a \rangle$ индекса j -го признака в $\Phi(a)$.

«Элементы каждого множества $S(\langle j, i \rangle), \langle j, i \rangle \in \sum \Phi(c)$ изначально упорядочены по возрастанию id^a , а значение j фиксировано для всех элементов в каждом из множеств $S(\langle j, i \rangle)$. Введём отношение линейного порядка Π_2 для применения операции объединения (слияния) линейно-упорядоченных множеств $S(\langle j, i \rangle)$: тройка $\langle id^a, j, \varphi_j^a \rangle$ предшествует тройке $\langle id^b, j', \varphi_{j'}^b \rangle$ тогда и только тогда, когда:

1. Выполнено неравенство $id^a < id^b$.
2. Выполнено равенство $id^a = id^b$, но при этом $j < j'$.

Операция слияния множеств $S(\langle j, i \rangle)$, полученных из ИИ для $\forall \langle j, i \rangle \in \Sigma \Phi(c)$, даёт линейно-упорядоченное множество $\tilde{S}(c)$, содержащее сведения о потенциально похожих изображениях-кандидатах в ИИ.

Оценка сходства эталонного изображения и изображений-кандидатов реализована за один проход по множеству $\tilde{S}(c)$, однако оценка перестаёт удовлетворять свойству симметричности и представляет собой псевдометрику. Для метрики Минковского (расстояния Хэмминга) оценка сходства эталонного изображения c и найденного в ИИ изображения a , принимает вид» [16]

$$\rho_M(c, a) = 1 - \frac{\sum_{j=1}^N g(\varphi_j^a, \varphi_j^c)}{2}, \quad (7.5)$$

$$\text{где } g(\varphi_j^a, \varphi_j^c) = \begin{cases} |\varphi_j^a - \varphi_j^c|, & \text{если } \langle id^a, j, \varphi_j^a \rangle \in \tilde{S}(c) \\ \max_{[x_i; y_i] \in \Delta} \{y_i | \varphi_j^a \in [x_i; y_i]\} - \min_{[x_i; y_i] \in \Delta} \{x_i | \varphi_j^a \in [x_i; y_i]\}, & \text{иначе} \end{cases}$$

«Свойство симметрии в формуле (7.5) отсутствует из-за вида функции $g(\varphi_j^a, \varphi_j^c)$: ввиду применения ИИ на этапе поиска имеется информация только по таким гистограммным словам и только для таких значений j , что значения φ_j^a и φ_j^c попали в один и тот же отрезок $[x_i; y_i]: \exists i \in \mathbb{N}, i \leq |\Delta|: [x_i; y_i] \in \Delta \& \varphi_j^a \in [x_i; y_i] \& \varphi_j^c \in [x_i; y_i]$. В противном случае делается предположение, что модуль разности $|\varphi_j^a - \varphi_j^c|$ не превосходит длины интервала, покрывающего отрезки, в которые попадает значение φ_j^c при квантовании (Рис. 7.1)» [16].

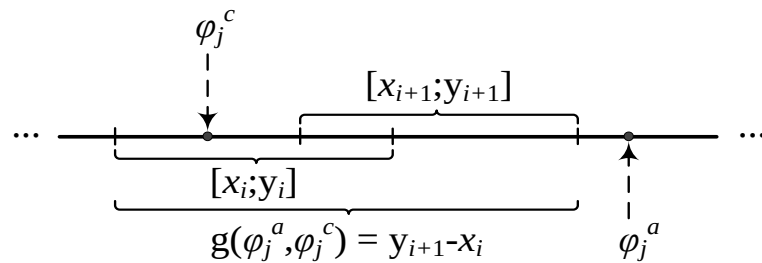


Рисунок 7.1 – Оценка разности $|\varphi_j^a - \varphi_j^c|$ на основе длины интервала, покрывающего отрезки при квантовании

7.3 Решение проблемы «замыкания цикла» на основе метода корзина слов и инвертированного индекса

Для отслеживания повторного прохождения мобильной платформой одного и того же места предполагается проводить на каждом шаге сравнение положения мобильной платформы с предыдущими шагами и в случае совпадения данных проводить корректировку движения и уточнение трехмерной карты окружающей среды [227-229]. В работе взято за основу и улучшено одно из известных решений этой проблемы: корзина слов (Bags of Binary Words) [230-235].

В корзине слов каждому RGB-D кадру (изображению) ставится в соответствие двумерный дескриптор, в работе используется предложенный в главе 3 двумерный дескриптор ДЛ ГНГ [19]. Информация о свойствах изображений храниться в «словаре», который организован в виде дерева. Каждый узел дерева содержит гистограммные слова, которые хранят информацию о весах слов, которые показывают насколько часто гистограммные слова встречались в обучающей выборке. Веса в дереве гистограммных слов представлены в виде вероятностей. Для того чтобы создать гистограммное слово в дереве, необходимо произвести поиск визуальных особенностей изображения на основе данных из обучающей выборки. Каждый следующий уровень дерева слов можно получить, повторяя описанную операцию с дескриптором родительского узла дерева. Таким образом, при анализе новых данных, поступающих на вход дерева гистограммных слов, производится процедура быстрого разложения идентификатора текущего кадра в линейную комбинацию узлов дерева гистограммных слов. Для того чтобы свести к минимуму ошибку метрики, используется граф оптимизации, основанный на ограничениях между положениями узлов в графе.

Интересующее событие («замыкание цикла») происходит, когда мобильная платформа возвращается в позицию, посещенную ранее в прошлом. Для определения «замыкания цикла» достаточно сохранять на каждом шаге историю прошлых RGB-D кадров и проверять: совпадает ли текущий RGB-D кадр с

некоторым из предыдущих кадров. Если текущий RGB-D кадр является визуально подобным с другим кадром, замеченным в прошлом, то преобразование может быть вычислено между двумя этими кадрами, и новое ограничение на графе положений может быть взято из данного кадра. Данная последовательность действий может быть повторена несколько раз. С помощью новых полученных ограничений положений кумулятивная ошибка графа может быть значительно уменьшена, все положения в графе могут быть связаны подобным способом, если такая связь в действительности существует. Представим детальную схему алгоритма обнаружения «замыкания цикла» [33] (Рис. 7.2):

Шаг 1. Сделать запрос к базе данных предыдущих RGB-D кадров, чтобы найти и вернуть наиболее похожий кадр, используя корзину слов.

Шаг 2. Пропустить всех результирующих кандидатов, которые принадлежат 20-ти последним ключевым RGB-D кадрам.

Шаг 3. Вычислить сходство (визуальное подобие) текущего изображения с каждым полученным кандидатом на основе их векторов V_1 , V_2 и игнорировать тех кандидатов, сходство которых будет ниже определенного порога SC , устанавливаемого эмпирическим путем.

Шаг 4. Возвратиться к K кадрам (изображениям) с наибольшим сходством в качестве кандидатов на «замыкания цикла». Выбор K будет зависеть от того, сколько кандидатов могут быть дополнительно проверены путем проверки геометрической согласованности в течение ограниченного периода времени.

Шаг 5. Текущий кадр добавляется в базу данных для того, чтобы его можно было позже найти в качестве кандидата на «замыкание цикла».

На рисунке 7.3 представлен план помещения из эталонного набора данных Apartment базы данных ASL DataSet [182], которая содержит семантическую маркировку пространства.

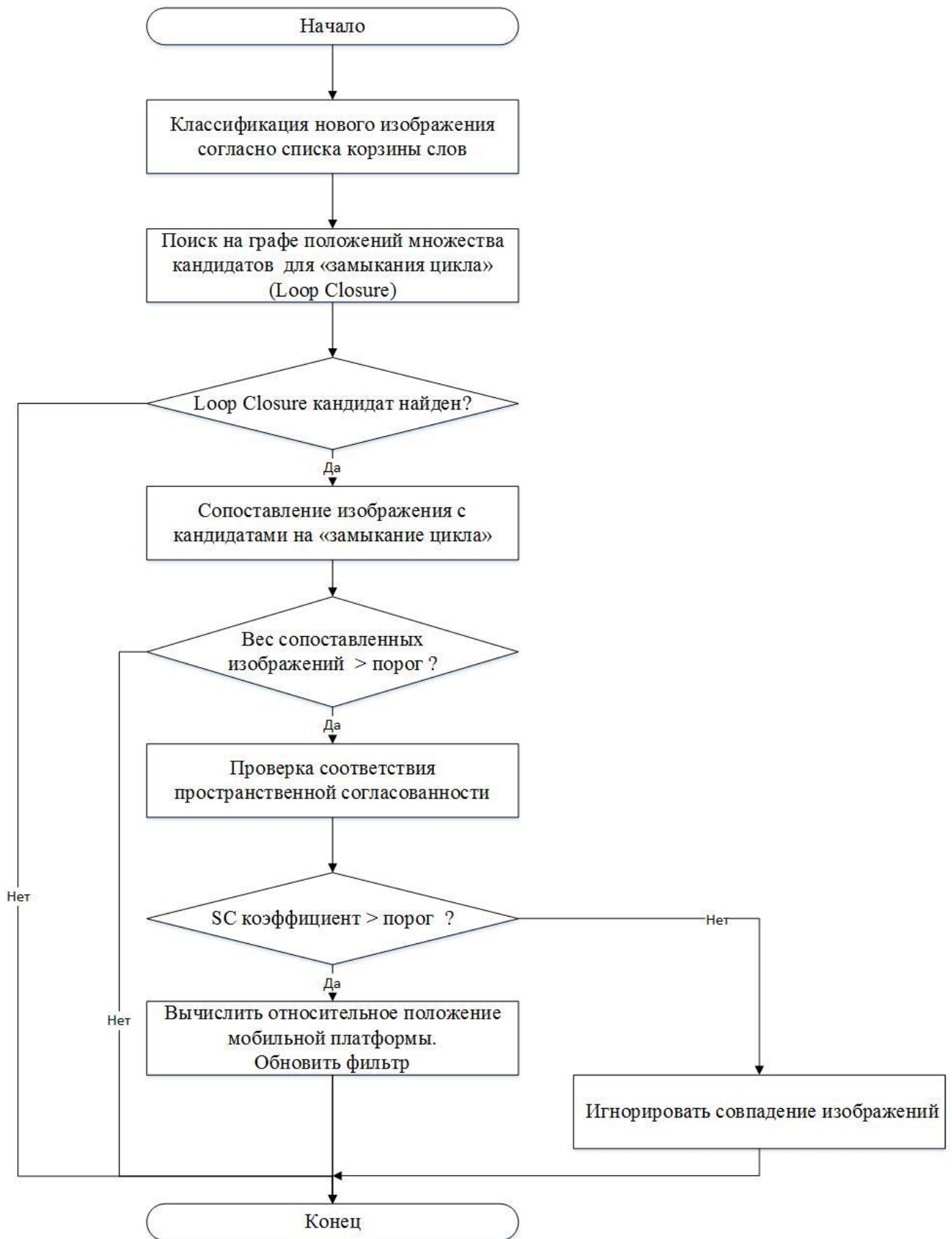


Рисунок 7.2 – Блок схема алгоритма обнаружения «замыкания цикла»

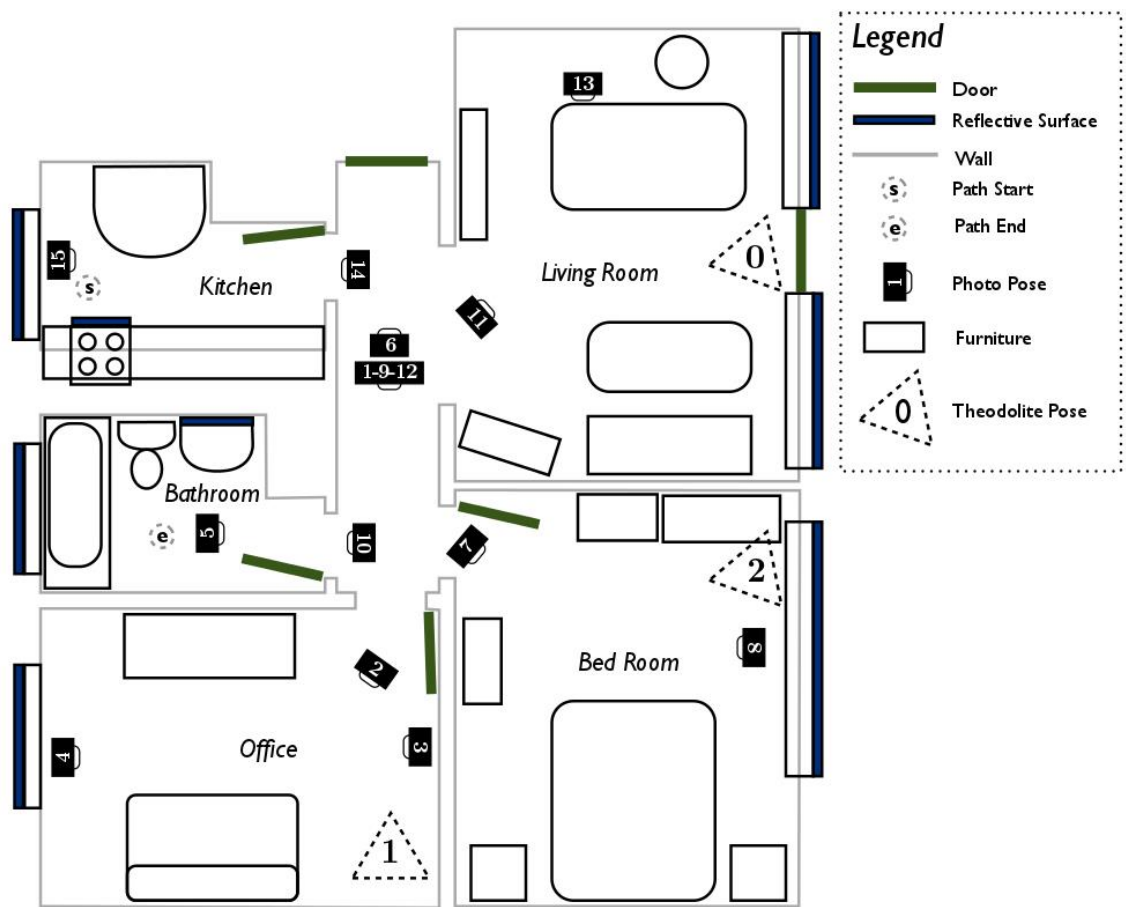
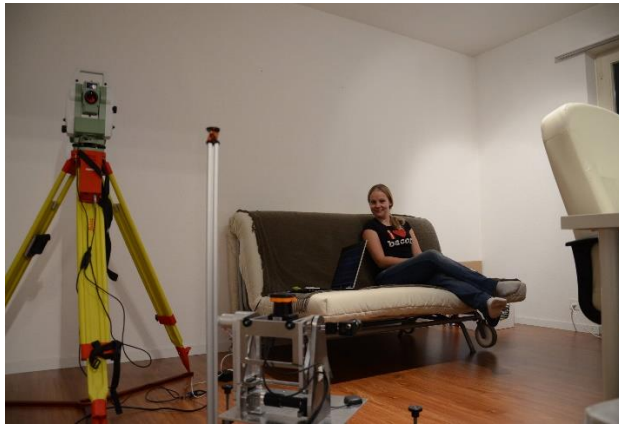


Рисунок 7.3 – Эталонный набор данных Apartment базы данных ASL DataSet.

План помещения

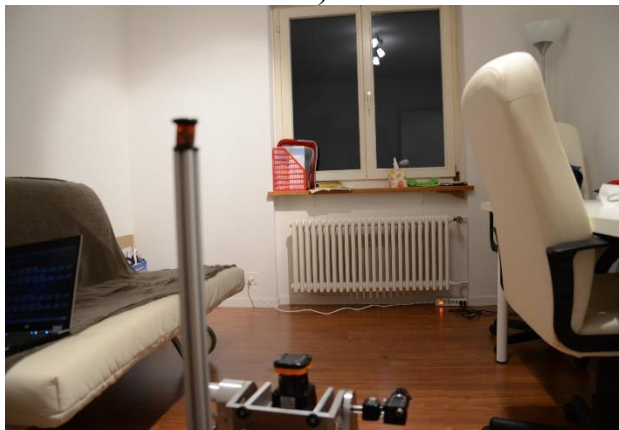
В работе для получения особых точек в кадрах и определения оптимального преобразования между последними ключевыми RGB-D кадрами используется предложенный двумерный дескриптор на основе дескриптора ДЛ ГНГ. RANSAC процедура применяется для более точного выравнивания кадров. Оптимизация графа положений приводит к глобально выровненному набору кадров RGB-D кадров. В более общем случае, метод слепой подстройки (bundle adjustment) одновременно оптимизирует граф положений и трехмерную плотную карту. На рисунке 7.4 представлен набор изображений (ключевые кадры) из эталонного набор данных Apartment базы данных ASL DataSet с обнаруженными большими циклами.



а)



б)



в)



г)



д)



е)

Рисунок 7.4 – Эталонный набор данных Apartment из базы данных ASL DataSet.

Калиброванные изображения помещения из разных положений камеры

Цикл на трехмерной сцене из эталонного набора данных Apartment состоит из 906 ключевых RGB-D кадров, общая длина траектории мобильной платформы на сцене более 84 м (Рис. 7.5). Для удобства представления данных наложим

трехмерную карту сцены на двумерный план сцены, при этом ряд точек, соответствующих структурным элементам сцены, например, множества точек для пола и потолка помещения удалим из представления.

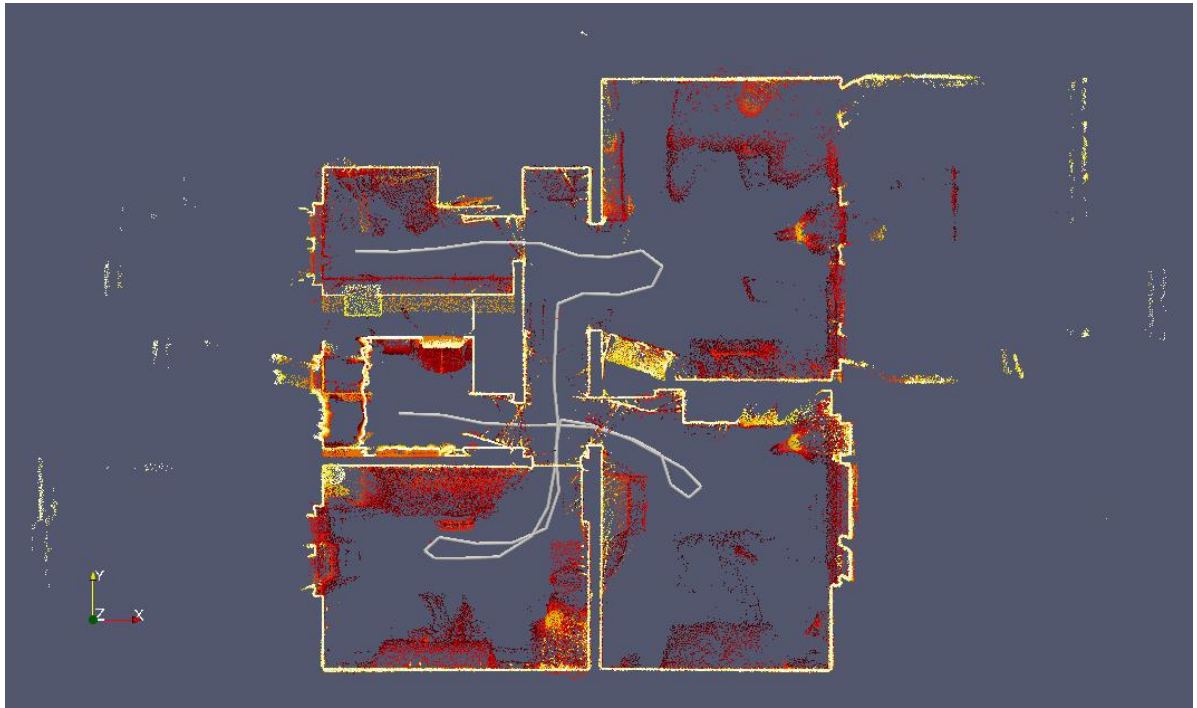


Рисунок 7.5 – Эталонный набор данных Apartment из базы данных ASL DataSet.

2D карта помещения с обнаруженными циклами

На Рис 7.6 представлены результаты построения траектории движения мобильной платформы для набора данных Apartment на основе семантических маркеров. Данная траектория движения содержит циклы, от которых необходимо избавиться. В результате компьютерного моделирования [42] было установлено: сходимость метода S-EKF SLAM значительно улучшается при использовании процедуры RANSAC, а сходимость метода определения динамического положения мобильной платформы улучшается за счет использования семантических ориентиров в расширенном фильтре Калмана. Метод S-EKF SLAM является результатом модификации известного метода EKF SLAM: Joint compatibility branch and bound (JCBB) [5, 7, 67]. Результаты компьютерного моделирования методов для коротких прогонов мобильной платформы (менее 100 метров) при отсутствии

процедуры «замыкания цикла» и при наличии процедуры «замыкания цикла» представлены в таблице 7.1 и таблице 7.2 соответственно. Произведено сравнение предложенного метода S-EKF SLAM с известными методами EKF SLAM [51, 53] и Graph-based SLAM (Graph SLAM) [57] при отсутствии одометрического датчика. Было установлено, что использование процедуры «замыкания циклов», в общем случае, позволяет улучшить точность локализации мобильной платформы на сцене для коротких циклов. Предложенный метод S-EKF SLAM имеет преимущества с точки зрения точности локализации по сравнению с известными методами SLAM.

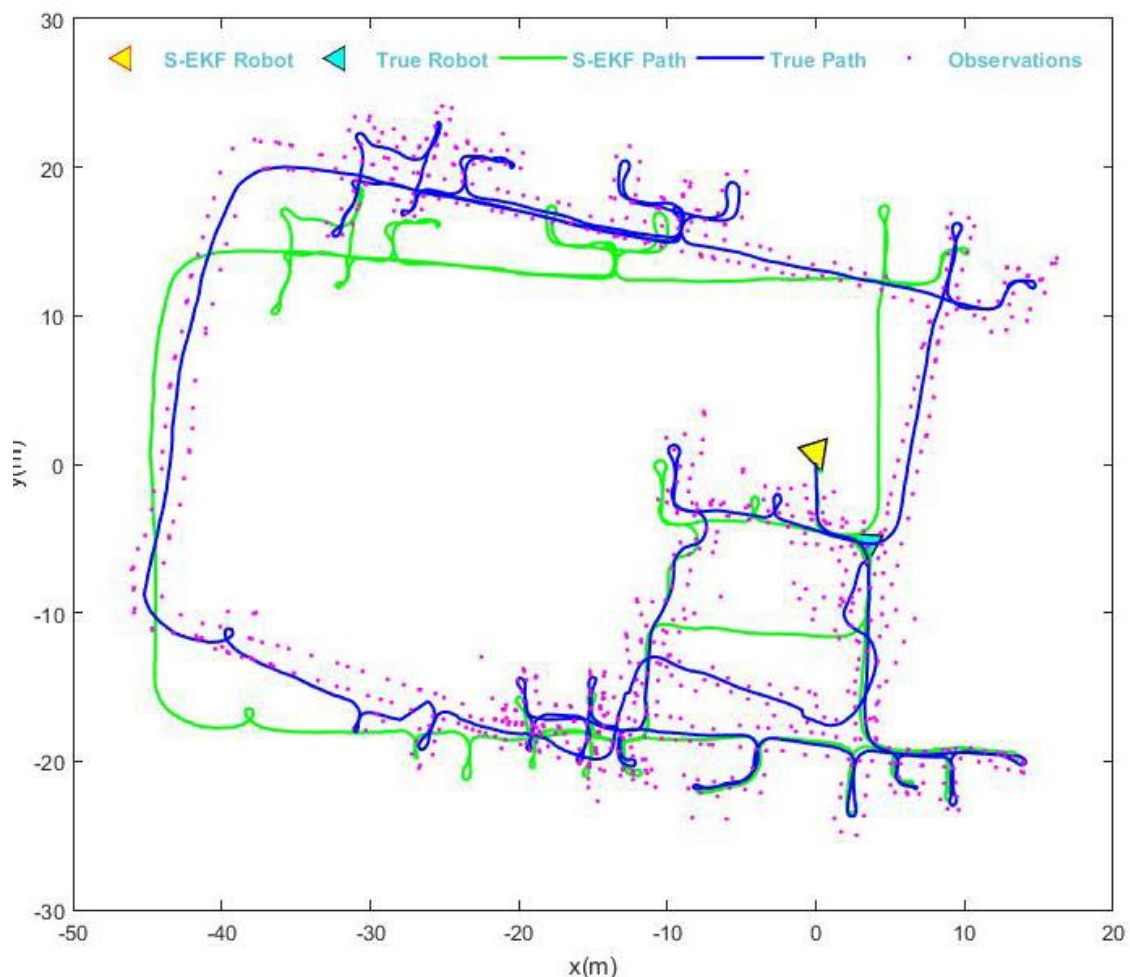


Рисунок 7.6 – Пример построения траектории мобильной платформы на основе S-EKF SLAM для набора данных Apartment

Аналогично проведем компьютерное моделирование для длинных циклов, в результате компьютерного моделирования был получен цикл длиной 452 метра,

содержащий 5123 кадра. Результаты компьютерного моделирования методов для длинных прогонов мобильной платформы при отсутствии процедуры «замыкания циклов» и при наличии процедуры «замыкания циклов» представлены в таблице 7.3 и таблице 7.4 соответственно.

Таблица 7.1 – Точность локализации методов SLAM без использования процедуры «замыкания цикла» (для коротких циклов)

Название метода	Среднее отклонение координат, мм	Средне-квадратичное отклонение координат, мм	Среднее отклонение ориентации, град	Средне-квадратичное отклонение ориентации, град
EKF SLAM	280,15	456,67	2,51	2,13
Graph SLAM	729,13	701,23	3,66	2,89
S-EKF SLAM	230,84	388,08	2,04	1,86

Таблица 7.2 – Точность локализации методов SLAM с использованием процедуры «замыкания цикла» (для коротких циклов)

Название метода	Среднее отклонение координат, мм	Средне-квадратичное отклонение координат, мм	Среднее отклонение ориентации, град	Средне-квадратичное отклонение ориентации, град
EKF SLAM	28,65	13,15	0,54	0,34
Graph SLAM	58,9	36,05	0,86	0,69
S-EKF SLAM	19,34	9,11	0,22	0,17

Было установлено, что использование процедуры «замыкания циклов», в общем случае, позволяет улучшить точность локализации мобильной платформы на сцене для длинных циклов. В целом же для коротких циклов точность

локализации методов EKF SLAM и S-EKF SLAM отличается, но не сильно (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). Таким образом, применение предложенного метода S-EKF SLAM позволяет получить преимущества в точности локализации для длинных циклов и позволяет избежать накопления ошибки со временем за счет распознавания ранее исследованных областей трехмерной сцены.

Таблица 7.3 – 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 7.4 – 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

7.4 Выводы по главе

1. «Представлен метод поиска изображений по визуальному подобию с применением ИИ гистограмм. Показано, что применение ИИ в задачах поиска визуально похожих изображений на основе сравнения гистограмм обеспечивает достаточную степень обобщения индексируемой информации и не приводит к потерям информации, что подтверждается результатами компьютерного моделирования. Применение ИИ позволяет получить существенный выигрыш с точки зрения производительности процедуры индексации» [16].

2. Предложен подход к решению проблемы «замыкания цикла» на основе известного алгоритма корзина слов с использованием визуальных характеристик неизвестной окружающей среды. Было установлено, что использование процедуры «замыкания циклов», в общем случае, позволяет существенно улучшить точность локализации мобильной платформы на сцене для длинных циклов, для коротких циклов применение процедуры имеет ограниченное влияние. Предложенный метод S-EKF SLAM имеет преимущества с точки зрения точности локализации по сравнению с известными методами SLAM.

3. Существующие методы сопоставления и регистрации данных, основанные на визуальной и одометрической информации, недостаточны для обеспечения реконструкции окружающего пространства. Комбинирование информации о глубине и цвете может обеспечить точное согласование и регистрацию RGB-D кадров, обнаружение «замыканий циклов». Возможно построить системы SLAM полного цикла, основанные только на недорогих RGB-D камерах.

4. Основные результаты по данной главе опубликованы в работах [15, 16, 19, 33, 42].

ГЛАВА 8 ПРАКТИЧЕСКИЕ ПРИЛОЖЕНИЯ

В данной главе рассмотрено практическое приложение комбинированной методологии SLAM при решении задачи спасения людей.

Разработанные в ходе выполнения диссертационного исследования методы, алгоритмы и программно-технические решения обеспечили создание аппаратно-программного комплекса, состоящего из группы мобильных платформ колёсного типа информационно-сервисного назначения и системы согласованного их функционирования в неизвестной окружающей среде. На разработанные методы и алгоритмы получены свидетельства о регистрации программы для ЭВМ [39-43] с целью внедрения предложенной методологии SLAM в коммерческие продукты компаний в гражданском секторе. Реализация задач диссертационного исследования выполнена посредством прорывных научных исследований и разработок мирового уровня в области информационных технологий. Все полученные научные результаты соответствует мировому уровню.

Основные направления использования результатов данной работы связаны с созданием систем управления автономными робототехническими комплексами:

- в различных областях наземного транспорта;
- в системах безопасности и спасения людей;
- при мониторинге за окружающим пространством в опасных или труднодоступных для человека объектах;
- в приложениях с дополненной реальностью.

В данной главе рассмотрены способы применения полученных результатов и их приложения. Важным и перспективным направлением результатов исследования является применение комбинированной методологии для решения проблемы одновременной навигации и построения карты в неизвестном пространстве для создания автоматических систем управления мобильной платформой в системах спасения людей, на которой установлены мультисенсорные биометрические датчики. 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.

Результаты, полученные в диссертационной работе, были апробированы в следующих организациях:

1. АО «Научно-производственное объединение Андроидная техника» – <https://npo-at.com>.
2. Югорский научно-исследовательский институт информационных технологий – <https://uriit.ru>.
3. ФГБОУ ВО «Челябинский государственный университет» в учебном процессе в рамках магистерской программы 02.04.02 «Фундаментальная информатика и информационные технологии» – <https://www.csu.ru>.
4. УРТЦ «Альфа-Интех» – <http://alphajet.ru>.
5. ООО «Научно-производственный центр Интеллектуальные транспортные системы» – <http://ittransport.ru>.

8.1 Применение комбинированной методологии решения проблемы одновременной навигации и картографирования в задаче спасения людей

В диссертационной работе рассматривается применение предлагаемой комбинированной методологии одновременной навигации и построения карты для создания систем спасения людей. Цель работы такой системы – идентификация в неизвестном окружающем пространстве личности и построение траектории движения мобильной платформы из текущей точки в пространстве в целевую точку, в которой находится искомый субъект. Системы спасения людей содержат три подсистемы:

1. Подсистема одновременной навигации и картографирования (SLAM).
2. Подсистема дистанционного некооперативного распознавания личности по лицу.
3. Подсистема планирования траектории мобильной платформы в неизвестном пространстве.

Первая подсистема позволяет построить точную трехмерную карту окружающей среды, при этом мобильная платформа с мультисенсорными датчиками осуществляет навигацию в неизвестной среде, пока не будет локализован искомый целевой субъект на сцене. В главах 3, 4, 5, 6 и 7 рассмотрены методы и алгоритмы, которые составляют в совокупности комбинированную методологию решения проблемы одновременной навигации и картографирования. На основе предложенной методологии реализована первая подсистема.

После того как личность на сцене была локализована используется вторая подсистема, которая выполняет процедуру дистанционного распознавания личности: если искомая личность распознана, то вычисляются координаты субъекта на исследуемой сцене. Далее используется третья подсистема и решается задача планирования траектории относительно исходной (текущей) позиции мобильной платформы и позиции, в которой находится искомый субъект на сцене.

В параграфе 5.4 был рассмотрен алгоритм планирования траектории мобильной платформы в неизвестном пространстве на основе метрико-топологического графа и семантических маркеров [180], который использован в подсистеме планирования траектории.

Если искомая личность не была распознана, то мобильная платформа продолжает движение на сцене. В соответствии с предложенной комбинированной методологией решения проблемы SLAM осуществляется итерационная процедура, в результате которой происходит реконструкция трехмерной карты для нового доступного фрагмента окружающего пространства и определение положения мобильной платформы на сцене. Далее для обновленной трехмерной карты окружающего пространства решается задача дистанционного некооперативного распознавания личности по лицу.

Для решения поставленной научно-технической задачи и решения задачи распознавания личности по лицу в диссертационной работе предложены следующие методы и схема:

- метод дистанционного распознавания личности по лицу с использованием автокодировщика, инвариантного к изменениям освещенности, позе человека, мимики лица;
- метод обратной индексации для поиска кандидатов по эталону для распознавания личности по лицу с использованием предложенного дескриптора ДЛ ГНГ;
- математическая схема комбинирования признаков и решений для мультимодального слежения за человеком внутри помещения с использованием нескольких камер.

Представим детальную графическую интерпретацию применения предлагаемой комбинированной методологии [11] для решения задачи спасения людей (Рис. 8.1) в виде итерационной процедуры, состоящей из следующей последовательности шагов:

Шаг 1. Получение данных с визуальных камер: камеры видимого диапазона Beward B2720 (2 шт., одна камера с широким углом обзора, другая камера с узким

углом обзора); RGB-D камера Kinect 2.0 (1 шт.); времяпролетная камера (ToF) MESA SR 4000 (1 шт.).

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 о визуально связанных характеристиках на изображении с использованием предложенного в работе двумерного дескриптора на основе локальных гистограмм ориентированных градиентов, вычисляемых в круглом скользящем окне (ДЛ ГНГ).

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

Шаг 6. Прогнозирование положения мобильной платформы (Husky UGV или Odyssey 6 Robotics) на основе расширенного фильтра Калмана (EKF).

Шаг 7. Сопоставление особых точек на двумерных и трехмерных данных в последовательности RGB-D кадров. Двумерные данные (изображения) сопоставляются с использованием предложенного в работе алгоритма сопоставления данных на основе дескриптора (ДЛ ГНГ). Трехмерные данные (облака точек) сопоставляются с использованием метода ближайших соседей.

Шаг 8. Реконструкция сцены на основе комбинированного метода ICP с метрикой точка-точка и метрикой точка-плоскость для группы аффинных и 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.

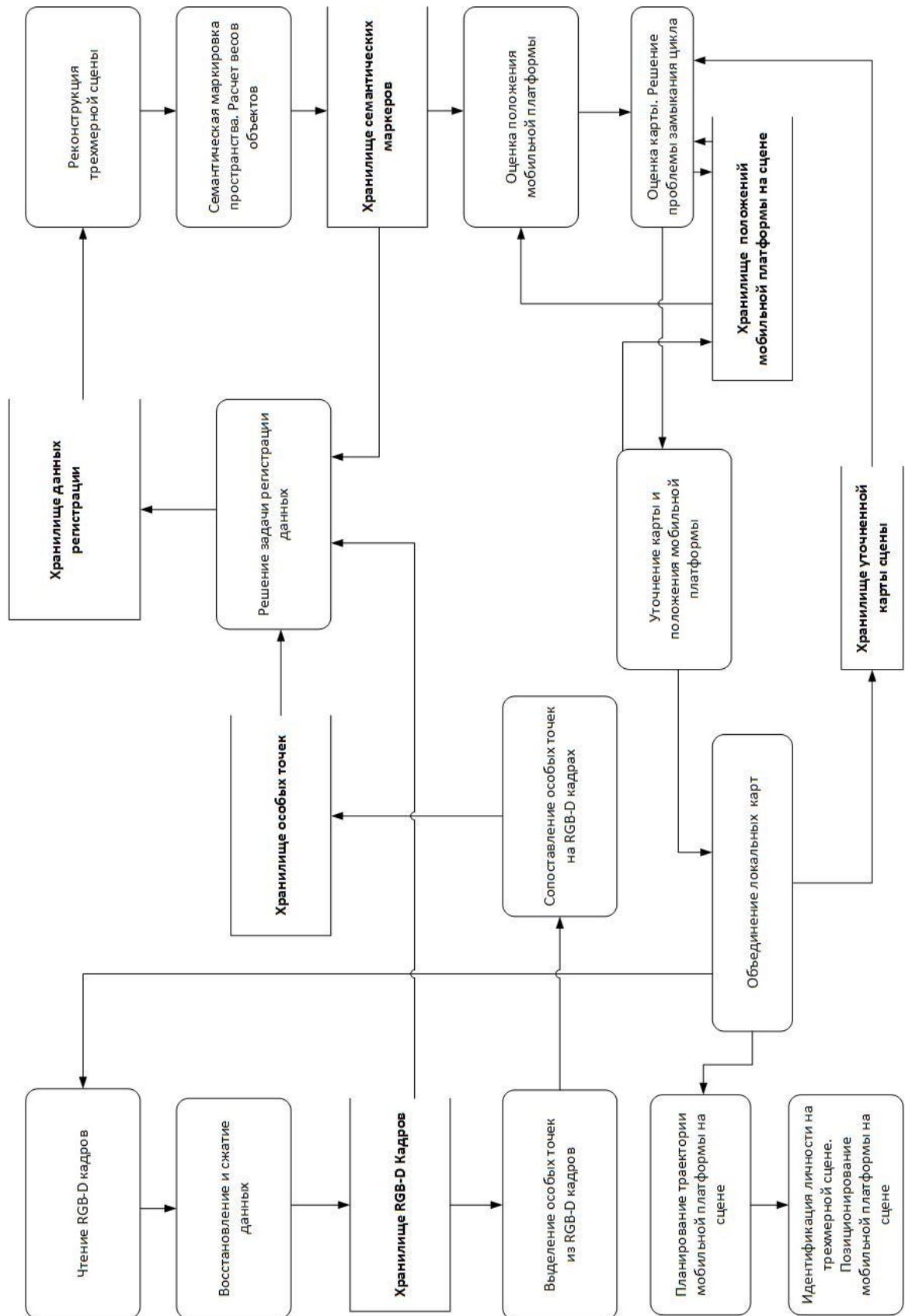


Рисунок 8.1 – Применение комбинированной методологии SLAM к задаче спасения людей

Шаг 11. Решение проблемы «замыкания цикла» с использованием предложенного дескриптора (ДЛ ГНГ), 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 мобильной платформы.

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

Шаг 14. Идентификация исходной и целевой позиции мобильной платформы на сцене.

Шаг 15. Идентификация личности на сцене. Сегментация области лица.

Шаг 16. Поиск изображений-кандидатов в коллекции изображений на основе метода обратной индексации для поиска лиц по эталону.

Шаг 17. Распознавание лица человека с использованием метода глубокого машинного обучения на основе автокодировщика.

Шаг 18. Решение задачи планирования траектории на основе метода HGA*.

В результате выполнения диссертационного исследования были получены следующие научно-технические результаты, на которые были получены свидетельства о регистрации интеллектуальной собственности:

1. Аналитический обзор современного состояния поставленной научно-технической проблемы.

2. Обоснование выбранного направления исследования и способов решения поставленных проблем по созданию комбинированной методологии одновременной навигации и картографирования.

3. Отчёт о патентных исследованиях в соответствии с ГОСТ Р 15.011-96 применительно к теме исследования НИР.

4. Анализ тенденции развития и перспектив применения автономных мобильных платформ для решения задач их согласованного функционирования в неизвестной окружающей среде.

5. Анализ существующих алгоритмов калибровки и коррекции данных, получаемых от RGB-D камер.
6. Концепция использования группы мобильных платформ.
7. Методы навигации мобильных платформ с использованием построенных трёхмерных карт окружающей среды.
8. Методы регистрации изображений относительно облака точек в трёхмерном пространстве на основе итеративного алгоритма ближайших точек.
9. Методы согласованного и оптимизированного планирования траекторий мобильных платформ с использованием построенной трёхмерной карты окружающего пространства с учётом возможных статических и динамических препятствий.
10. Методы управления пространственно-распределённой группой мобильных платформ.
11. Экспериментальный комплекс, состоящий из группы мобильных платформ, созданный с целью их одновременной навигации с использованием разработанной технологии технического зрения.
12. Методические пособия и рекомендации по использованию произведённой группы мобильных платформ.
13. Конструкторская и программная документация на экспериментальный комплекс, состоящий из группы мобильных платформ в составе трёх колёсных сервисно-информационных платформ Husky UGV (2 шт.) и Odyssey 6 Robotics (1 шт.).
14. Техническое задание на создание экспериментального комплекса, состоящего из группы мобильных платформ.

8.2 Метод дистанционного распознавания личности по лицу с использованием автокодировщика

Для разработки системы распознавания личности (субъекта) по изображениям лиц с применением нейронных сетей глубокого обучения с качеством распознавания близким к возможностям человека в работе разработаны новые математические методы и алгоритмы, включающие в себя обработку изображений, выбор признаков и классификацию. Выбор признаков проводится автоматически с использованием техник глубокого обучения. Главная проблема, которую необходимо решить – это использование глубокого обучения при относительно небольшом объеме данных [236]. Для увеличения объема данных в работе используются генеративные состязательные сети (Generative Adversarial Networks) для синтеза новых изображений на основе уже существующей базы данных [237]. Состязательная сеть позволяет идентифицировать лицо человека вне зависимости от мимики и угла поворота головы и позволяет извлекать основные черты лица на изображениях, даже при внесении значительных помех. Кроме того, геометрические преобразования над существующими изображениями также помогают искусственно увеличить обучающую выборку. Таким образом, для разработки системы распознавания личности по лицу необходимо решить следующие задачи:

1. Разработать алгоритм нормализации изображений.
2. Разработать метод и алгоритм выбора признаков, используя методы глубокого обучения, например, автокодировщики [238, 239] для обеспечения высокой информативности с низкой внутриклассовой и высокой межклассовой вариативностью.
3. Разработка алгоритмов соответствия изображений и архитектуры глубокой сверточной нейронной сетей для классификации изображений.

Рассмотрим современное состояние проблемы. Разрабатываемая система обработки, анализа и классификации субъектов по изображениям их лиц обеспечивает качество распознавания близкое к возможностям человека.

Рассмотрим метод распознавания и выбора признаков человеческих лиц с использованием техник глубокого обучения. Эта задача является одной из составных частей в создании системы спасения людей. Для задачи распознавания личности по лицу был разработан автокодировщик (autoencoder), представляющий собой глубокую искусственную нейронную сверточную сеть ResNet с 29 слоями [27]. Для ее настройки применялась техника глубокого остаточного обучения. При использовании искусственных нейронных сетей основной задачей является их обучение. В реальности не всегда удастся найти достаточное количество размеченных данных. При этом, чем сложнее архитектура нейронной сети, тем больший объем данных необходим для обучения сети, чтобы избежать переобучения. Альтернативой обучения с учителем является обучение без учителя, примером которого служат автокодировщики [238]. Под автокодировщиком понимается нейронная сеть, реализующая отображение входа сети на себя. Но при отображении внутри автокодировщика происходит сужение размерности входных данных. Для того чтобы получить на выходе автокодировщика объект близкий к входным данным, необходимо выделить основные признаки входа. Применительно к задаче распознавания лиц автокодировщики позволяют выделять информативные черты лица [240]. Значительным преимуществом автокодировщиков является отсутствие необходимости размечать входные данные, что упрощает процесс построения нейронной сети [238, 240]. Также частично решается проблема пополнения базы изображений, так как после обучения автокодировщика можно в качестве входных данных использовать изображения, полученные автокодировщиком [239].

Обобщенный автокодировщик состоит из двух частей: кодировщика и декодировщика. Автокодировщик карт ввода $x_i \in R^{d_x}$ к уменьшенному скрытому представлению $y_i \in R^{d_y}$ задается функцией g следующего вида

$$y_i = g(W'x_i), \quad (8.1)$$

где g может быть как функцией тождества для линейной проекции, так и сигмной функцией вида $\frac{1}{1+e^{-Wx}}$ для нелинейного отображения.

Параметр W' есть весовая матрица размером $d_y \times d_x$. В работе игнорируются условия смещения нейронной сети для упрощения представления структуры автокодировщиков. Декодировщик реконструирует $x_i \in R^{d_x}$ из скрытого представления (слоя) y_i

$$x'_i = f(W'y_i). \quad (8.2)$$

Общая ошибка реконструкции E для n образцов может быть вычислена следующим образом

$$E(W, W') = \sum_{i=1}^n e_i(W, W') = \sum_{i=1}^n \sum_{j \in \Omega_i} s_{ij} L(x_j, x'_i), \quad (8.3)$$

где L – ошибка реконструкции, $\Omega_i = \{j, k, \dots\}$ – набор индексированных образцов с определенными весами $S_i = \{s_{ij}, s_{ik}, \dots\}$ для x_i .

Обобщенный автоматический автокодировщик обучает параметры (W, W') , минимизируя общую ошибку реконструкции E .

Представим алгоритм итеративной процедуры обучения для обобщенного автокодировщика (Рис. 8.2):

Входная обучающая выборка $\{x_i\}_1^n$. Параметры: $\Theta = (W, W')$. Система обозначений: Ω_i – реконструированное множество для x_i , S_i – набор реконструированных весов x_i , $\{y_i\}_1^n$ – скрытое представление.

Шаг 1. Вычислим реконструированные веса S_i из $\{x_i\}_1^n$ и определим реконструируемый набор Ω_i , например, используя метод ближайшего соседа (k-nearest neighbor).

Шаг 2. Минимизируем общую ошибку реконструкции E из выражения (8.3), используя стохастический метод градиентного спуска и обновим значения весов Θ для t шагов.

Шаг 3. Вычислим скрытое представление (слой) $\{y_i\}_1^n$, и обновим S_i и Ω_i из $\{y_i\}_1^n$ соответственно.

Шаг 4. Повторяем Шаги 2 и 3 до тех пор, пока алгоритм не сойдется.

При использовании линейных вложенных графов можно предположить, что низкоразмерные представления могут быть получены на основе линейной проекции $y = X^T w$, где w есть проекция вектора $X = [x_1, x_2, \dots, x_n]$. Целевая функция LGE может быть получена как

$$w^* = \underset{w^T X B X^T w = c \text{ or } w^T w \neq c}{\operatorname{argmin}} \sum_{ij} \|w^T x_i - w^T x_j\|^2 s_{ij}. \quad (8.4)$$

Генеративная состязательная сеть [25] стремится использовать пространственную информацию между пикселями изображения. Для простоты предполагается, что изображение в градациях серого цвета определяется следующей функцией

$$I: \{1, \dots, n_1\} \times \{1, \dots, n_2\} \rightarrow W \subseteq R, (i, j) \mapsto I_{i,j}. \quad (8.5)$$

Изображение I может быть представлено массивом размером $n_1 \times n_2$. Учитывая фильтр, дискретная свертка изображения I с фильтром K задается так, что изображение I может быть представлено как массив размера $n_1 \times n_2$. Если учесть $K \in R^{2h_1+1 \times 2h_2+1}$, то дискретная свертка изображения I с фильтром K может быть задана следующей функцией

$$(I * K)_{r,s} := \sum_{u=-h_1}^{h_1} \sum_{v=-h_2}^{h_2} K_{u,v} I_{r+u, s+v}, \quad (8.6)$$

где фильтр K задается следующей матрицей

$$K = \begin{pmatrix} K_{-h_1, -h_2} & \dots & K_{-h_1, h_2} \\ \dots & K_{0,0} & \dots \\ K_{h_1, -h_2} & \dots & K_{h_1, h_2} \end{pmatrix}. \quad (8.7)$$

Поведение данной операции по отношению к границам изображения должно быть определено правильно. Для сглаживания будем использовать фильтр в виде дискретного гауссовского фильтра $K_{G(\sigma)}$, который может быть задан в виде следующей формулы

$$(K_{G(\sigma)})_{r,s} = \frac{1}{\sqrt{2\pi}\delta^2} \exp\left(\frac{r^2 + s^2}{2\delta^2}\right), \quad (8.8)$$

где σ – стандартное отклонение гауссовского распределения.

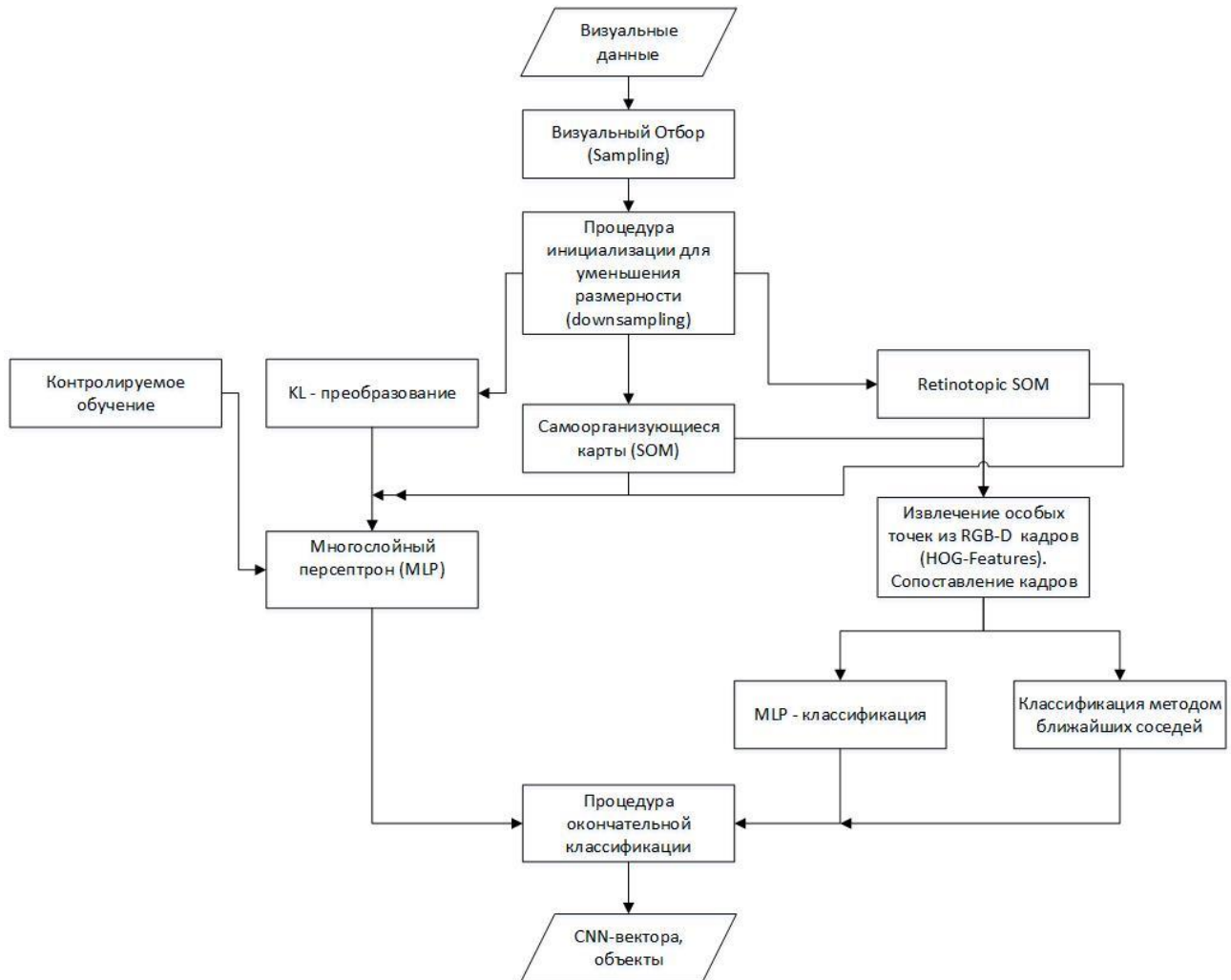


Рисунок 8.2 – Блок-диаграмма метода выбора особых точек на основе автокодировщиков

Продemonстрируем эффективность предложенного метода для выбора признаков с использованием методов глубокого обучения на основе автокодировщиков при моделировании распределения структурированных выходных переменных. Для подтверждения эффективности предложенного метода создана искусственная экспериментальная настройка для структурированного прогнозирования выходных данных с использованием базы данных LFW (Labeled Faces in the Wild) [240]. Проведем сравнение трех нейронных сетей (НС): эталонная

базовая модель нейронной сети (Baseline), CVAE модель на основе подхода Монте-Карло и предложенная модифицированная модель на основе CVAE. Для компьютерного моделирования произведено обучение данных моделей и произведено сравнение их производительности. Результаты представлены в таблице 8.1.

Таблица 8.1 – Точность и производительность автокодировщика для выбора особых точек

Модель НС	1 квадрант		2 квадрант		3 квадрант	
	утвер- ждение	экспе- римент	утвер- ждение	экспе- римент	утвер- ждение	экспе- римент
Точность, значение ошибки						
Baseline	100	99.44	57.63	57.98	20.01	19.17
CVAE (Монте-Карло)	57.96	57.46	39.45	39.04	17.14	17.11
Предложенный CVAE	55.04	54.88	32.48	32.15	15.13	15.04
Вычислительная сложность, сек						
Baseline	1.6		1.53		1.46	
CVAE (Монте-Карло)	1.22		1.11		0.89	
Предложенный CVAE	0.98		0.51		0.36	

Обучение разработанного автокодировщика проводилось на 3 млн. изображений лиц из базы данных LFW. Результаты компьютерного моделирования показали, что разработанный автокодировщик позволяет хорошо извлекать характерные черты лица. Другой важной проблемой при распознавании изображений является их зашумленность. При наличии значительных помех автокодировщик не способен адекватно выделять признаки. Для решения этой задачи часто применяется алгоритм регуляризации двумерной полной вариации, позволяющий подавлять шум на изображениях. Вычисление экстремальной функции задачи регуляризации основано на редукции двумерной задачи регуляризации к одномерной. Данный алгоритм позволяет в комбинации с нейросетевым автокодировщиком выделять характерные черты лица, даже в случае наличия существенных помех.

8.3 Метод обратной индексации для поиска лиц по эталону в задаче спасения людей

Рассмотрим применение предложенного в параграфе 7.2 метода индексации данных на основе инвертированного индекса (ИИ) применительно к задаче быстрого нахождения изображений-кандидатов в эталонных базах данных лиц человека [14, 21]. Пусть имеется набор изображений, содержащий кадры данных с неsegmentированными лицами человека $I = \{a_i\}_{i=1}^M$. Для решения поставленной задачи: создадим ИИ данного набора данных и разработаем метод поиска изображений лиц человека по визуальному подобию. Обозначим исходное изображение, для которого осуществляется поиск изображений-кандидатов, как d .

Рассмотрим основные этапы предлагаемого метода поиска лиц по эталону:

1. Распознавание лица и выделение области, содержащий его. Для распознавания лица в работе используется известный метод Виолы-Джонса [243, 244]. После идентификации личности в кадре происходит сегментация области на изображении, включающей лицо человека, и создается новое изображение f для выделенного сегмента.

2. Процедура выравнивания лица относительно центра изображения. На выделенном сегменте f определим расположение области глаз на основе алгоритма Виолы-Джонса. Далее вычислим отдельно координаты центра для правого и левого глаза. На основе координат для центров глаз определим направление поворота изображения относительно двумерной сцены в соответствующем кадре

$$direction = \langle Reye_x - Leye_x, Reye_y - Leye_y \rangle, \quad (8.9)$$

где $direction$ – направление поворота, $Reye_x$ – x -координата центра правого глаза, $Reye_y$ – y -координата центра правого глаза, $Leye_x$ – x -координата центра левого глаза, $Leye_y$ – y -координата центра левого глаза.

Для удобства положим, что центр поворота изображения совпадает с координатами центра левого глаза ($Leye_x, Leye_y$). Учитывая этот факт, определим значение угла поворота (в радианах):

$$angle = \arctan(direction_x, direction_y). \quad (8.10)$$

Далее для каждой точки изображения f выполним выравнивание лица, используя информацию о координатах центра изображения. Тогда новые координаты на изображении можно определить, используя следующую систему уравнений:

$$\begin{cases} x'_i = a_0 + a_1 x_i + a_2 y_i \\ y'_i = b_0 + b_1 x_i + b_2 y_i \end{cases}, \quad (8.11)$$

где (x_i, y_i) – координаты i -ой точки изображения-кандидата f ;

(x'_i, y'_i) – координаты i -ой точки на выравненном изображении f' ; $a_i, b_i, i=\overline{0,2}$;

$$a_0 = \cos(angle), a_1 = \sin(angle), a_2 = (1 - a_0)Leye_x - a_1 Leye_y;$$

$$b_0 = -\sin(angle), b_1 = \cos(angle), b_2 = (1 - b_1)Leye_y - b_0 Leye_x.$$

3. Сегментация зон на лице. На изображении f' с использованием метода Виолы-Джонса выделим сегменты лица $p_i, i = 1, 4$, которые дополнительно должны быть приведены к одинаковому размеру (64x64). Сегментацию изображения проведем на следующие части: область носа, область вокруг рта, области вокруг левого глаза, область вокруг правого глаза. Данные сегменты далее обрабатываются по отдельности.

4. Выделение ключевых признаков. Для каждого сегмента p_i определим ключевые признаки с использованием дескриптора, предложенного в работе Далала Н. [245]. Для каждого сегмента p_i на изображении-кандидате a построим дескриптор φ^i :

$$\varphi^i(a) = \langle \varphi_1^i(a), \dots, \varphi_N^i(a) \rangle, \quad (8.12)$$

где $i=\overline{1,4}$; N – мощность множества признаков.

Построим дескриптор $\Phi(a)$ на основе слияния отдельных векторов признаков $\varphi^i, i = \overline{1,4}$

$$\begin{aligned} \Phi(a) = \langle \varphi^1(a), \varphi^2(a), \varphi^3(a), \varphi^4(a) \rangle &= \langle \varphi_1^1(a), \dots, \varphi_N^1(a) \rangle, \\ &\langle \varphi_1^2(a), \dots, \varphi_N^2(a) \rangle, \langle \varphi_1^3(a), \dots, \varphi_N^3(a) \rangle, \langle \varphi_1^4(a), \dots, \varphi_N^4(a) \rangle. \end{aligned} \quad (8.13)$$

Размерность вектора при этом

$$|\Phi(a)| = M = 4 * N. \quad (8.14)$$

5. Квантификация и индексация. Для организации эффективного поиска в эталонных базах изображений, которые содержат большие объемы данных, используют методы индексации на основе ИИ [246]. Для построения ИИ необходимо всем компонентам вектора признаков $\Phi(a)$ сопоставить «гистограммные слова» при помощи процедуры квантификации [15]. В исследовании для квантификации применены нормализованные дескрипторы на основе гистограмм ориентированных градиентов, поэтому все элементы дескриптора $\Phi(a)$ размещены в диапазоне [0;1].

6. Поиск изображений-кандидатов. «На стадии распознавания лица на изображении d нужно построить вектор признаков $\Phi(d)$ и затем построить прямой индекс $DI(d)$. Гистограммные слова $\langle p, c, i \rangle$ построенного прямого индекса будут запросами к инвертированному индексу коллекции. Оценка сходства эталонного изображения и изображений-кандидатов может быть выполнена с помощью однопроходного алгоритма по множеству $\tilde{S}(d)$ на основе псевдометрики» [14]

$$\rho(d, a) = 1 - \sum_{j=1}^N \sum_{p=1}^4 g(\varphi_j^p(d), \varphi_j^p(a)) / M. \quad (8.15)$$

Представим результаты компьютерного моделирования. В работе была разработана подсистема для распознавания личности по лицу на основе библиотеки Open Source Computer Vision Library (OpenCV). В качестве программного обеспечения для реализации ИИ была выбрана библиотека Light Memory Database (LMDB). Входные данные для данной библиотеки: покрытие отрезка [0;1], длины $\alpha=0.125$, перекрытие $\beta=\frac{\alpha}{2}=0.0625$. В качестве эталонных баз данных, содержащих изображения лиц человека, в работе были использованы следующие БД: MegaFace Benchmark [247] и Labeled Faces in The Wild [248]. Рассмотрим процесс обработки изображения для представленного метода поиска изображений-кандидатов на примере первой базы данных в виде следующей последовательности шагов:

Шаг 1. Выберем из БД LFW файл с изображением "Aaron_Eckhart_0001" (Рис. 8.3).

Шаг 2. Выполним сегментацию области лица человека на данном изображении.

Шаг 3. Выполним выравнивание лица человека.

Шаг 4. Выделим ключевые признаки на изображении и получим набор изображений с частями лица $p_i, i=\overline{1,4}$ (Рис. 8.4).

Шаг 5. Определим ключевые признаки для каждого изображения p_i из набора с использованием предложенного дескриптора ДЛ ГНГ.

Шаг 6. Выполним процедуру квантификации. Результаты данного шага сохраняются в структуре ИИ.



Рисунок 8.3 – Пример изображения из коллекции LFW-файл
Aaron_Eckhart_0001.jpg

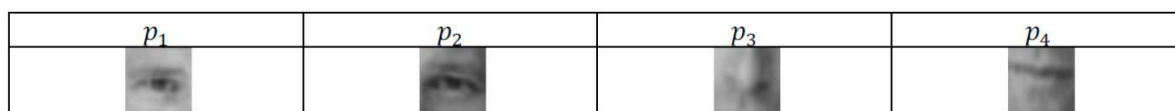


Рисунок 8.4 – Выделенные зоны лица человека

Предлагаемый метод применяется для предварительного отбора изображений-кандидатов, которые далее должны будут проходить заключительную обработку для более точного определения подобия [14]. Поэтому наиболее важными для системы являются показатели полноты. Для изображения a вычислим полноту (recall) по формуле:

$$recall_a = \frac{TP_a}{TP_a + FN_a}, \quad (8.16)$$

где TP_a – правильно найденные похожие изображения, FN_a – неправильно отвергнутые системой изображения.

«Для расчета эффективности работы системы на всей коллекции необходимо вычислить среднюю полноту для ее изображений. Для этого существует два подхода. Первый заключается в оценке микроусреднения

$$recall_{micro_av} = \frac{\sum_{i=1}^K TP_i}{\sum_{i=1}^K TP_i + \sum_{i=1}^K FN_i}, \quad (8.17)$$

где K – количество изображений в коллекции.

Альтернативной метрикой качества является макроусреднение

$$recall_{macro_av} = \frac{\sum_{i=1}^K recall_i}{K}. \quad (8.18)$$

В ходе проведения эксперимента варьировался порог сходства для оценки $\rho(d, a)$ из (8.15), таким способом, изменялось количество найденных схожих изображений (Табл. 8.2)» [14].

Таблица 8.2 – Макро- и микроусреднение значений полноты в зависимости от порога сходства

Порог сходства	$recall_{micro_av}$	$recall_{macro_av}$
0,2	0,998395	0,998668
0,205	0,997655	0,998274
0,215	0,994322	0,995657
0,23	0,984942	0,988131
0,2316	0,983584	0,987014
0,2332	0,982041	0,986106
0,2348	0,98056	0,984951
0,2364	0,978586	0,982961
0,2396	0,973525	0,979904
0,246	0,963157	0,970578
0,25	0,953468	0,965361
0,254	0,942854	0,958729
0,275	0,868428	0,900039
0,305	0,691558	0,788452
0,34	0,442236	0,635048

«Пример результата поиска сходных изображений по заданному эталонному изображению представлен на рисунке 8.5 и рисунке 8.6. Эксперимент показывает,

что предлагаемый подход помогает найти похожие изображения-кандидаты с высокой полнотой, когда порог сходства имеет адекватное значение для рассматриваемой коллекции» [14].



Рисунок 8.5 – Эталонное изображение из коллекции LFW-файл
"Angelina_Jolie_0014.jpg"



Рисунок 8.6 – Изображения-кандидаты для заданного эталонного изображения

8.4 Выводы по главе

1. Предложенный алгоритм распознавания лиц на основе автокодировщика при более высокой точности сопоставления топологических моделей в сравнении с известными аналогами свободен от многих существующих проблем некооперативных методов распознавания личности. Таким образом, применение данного алгоритма в системах безопасности и спасения людей позволит улучшить интегральную точность обнаружения субъектов в данных системах и осуществить слежение за объектами внутри помещения в зонах с высокой зашумленностью или в «слепых зонах».

2. Предложен эффективный метод поиска изображений-кандидатов на основе инвертированного индекса для решения задачи распознавания личности по лицу. Данный метод осуществляет поиск-изображений кандидатов в эталонных базах данных с высокой полнотой как с метрикой микроусреднения, так и с метрикой макроусреднения. Полученные изображения-кандидаты впоследствии подвергаются обработке и анализу в алгоритме, который использует сопоставление данных с эталонной двумерной моделью лица. Предложенный подход позволяет существенно повысить точность распознавания личности за счет использования комбинированной адаптивной динамической схемы. При этом процесс идентификации личности происходит в реальном масштабе времени.

3. В работе рассмотрено применение разработанной комбинированной методологии одновременной навигации и построения карты при создании системы спасения людей.

4. Основные результаты по данной главе опубликованы в работах [14, 17, 21, 22, 25, 27, 38, 43].

ЗАКЛЮЧЕНИЕ

В работе предложена комбинированная методология решения проблемы одновременной навигации и построения карты на основе слияния мультисенсорной динамической информации и информации о семантических характеристиках окружающей среды, пространственного совмещения последовательных кадров данных, обнаружения «замыкания цикла» и определения положения мобильной платформы на сцене. 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 может быть 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, инвариантных 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 ИСР 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 (EKF), 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. 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 методологии Visual SLAM и методологии EKF SLAM, 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.

Перспективы дальнейших исследований – дальнейшее развитие методологии одновременной навигации и картографирования связано с решением следующих задач:

- разработка методов реконструкции трехмерных данных для сцен, содержащих не ригидные объекты;
- разработка методов планирования траектории для сцен с динамическими объектами;
- разработка адаптивных методов слежения за трехмерными объектами на динамических сценах;
- развитие методов семантической маркировки неизвестного окружающего пространства с целью повышения точности классификации объектов на сцене;
- создание методов SLAM для навигации на сценах, искаженных помехами окружающей среды различной природы;
- исследование позиционной точности предложенного метода S-EKF SLAM;
- разработка принципов коллективного технического зрения группы мобильных платформ;
- разработка методов коллективного построения плотных динамических трехмерных карт окружающей среды с высокой точностью на основе предложенного метода S-EKF SLAM.

Словарь терминов

Мобильная платформа	: Объект, который движется в неизвестной окружающей среде, местоположение которого в ней требуется оценить.
Комбинированное решение вариационной задачи ICP	: Метод реконструкции окружающей среды (построения карты, картографирования) с использованием семантических и визуальных характеристик окружающей среды.
Комбинированная методология	: Методология решения проблемы SLAM на основе комбинирования визуальных и семантических характеристик неизвестной окружающей среды.
Навигация и определение положения мобильной платформы	: Метод решения задачи локализации мобильной платформы в неизвестной окружающей среде и идентификации ориентиров в каждый момент времени.
Ключевой кадр	: Система отчета, которая используется в SLAM для локализации в окружающей среде. Ключевой кадр содержит данные о глубине и цвете, полученные с камеры в каждый момент времени.
Замыкание цикла	: Ограничение между положениями мобильной платформы в графе положений. Ограничение определяется при возвращении платформы в положение, которое ранее было пройдено.
Облако точек	: Набор данных, состоящий из множества точек, различают плотные (получаемые с датчика) и разреженные облака точек. В работе исследуется

трехмерные облака точек в виде данных о глубине сцены.

- Граф положений : Представление траектории движения мобильной платформ в виде графа, в котором каждая вершина соответствует позиции платформы на сцене и каждое направленное ребро соответствует некоторому преобразованию (поворот, перенос или изменение масштаба) положения платформы на сцене.
- Оптимизация графа положений : Поиск решения линейной оптимизационной задачи, заданной ограничениями замыкания цикла в графе положений.
- Сцена / Окружающая среда/ Пространство : Физический (материальный) мир.
- Карта/ Карта сцены : Набор данных сцены, который соответствует исследованной мобильной платформой области окружающего пространства и может представлять облако точек или карту особых точек. Различают локальные карты сцены и глобальную карту сцены.
- Особые точки/ визуально связанные характеристики окружающей среды : Визуальные шаблоны или структуры, которые обладают высокой различительной способностью и могут быть использованы для сравнения и отслеживания (характерные точки).
- Корзина слов : Вектор, который содержит число вхождений объектов на изображении в древовидный словарь объектов. Объект соответствует множеству особых точек на изображении и выделяется из него при помощи ядра.
- Семантические маркеры (ориентиры) : Навигационные ориентиры на сцене, которые связаны с принадлежностью объекта к определенному классу предметной области.

Список сокращений и условных обозначений

АП	-	Аддитивная помеха
АШ	-	Аддитивный шум
ГНГ	-	Гистограммы направленных градиентов
ДЛ ГНГ (HOGs)	-	Дескриптор на основе гистограмм ориентированных градиентов
ДКП	-	Дискретное косинусное преобразование
ИИ	-	Инвертированный индекс
МНП	-	Неоднородная мультипликативная помеха
МП	-	Мультипликативная помеха
МТ-граф	-	Метрико-топологический граф
МФ	-	Медианный фильтр
НАП	-	Неоднородная аддитивная помеха
ЭК	-	Экспериментальный комплекс, состоящий из трех колесных мобильных платформ информационно-сервисного типа
3D сцена	-	Трехмерная сцена
3DEF	-	Метод трехмерной семантической сегментации на основе запутанных лесов (3D Entangled Forest)
EKF	-	Расширенный фильтр Калмана (Extended Kalman Filter)
FAST SLAM	-	Particle Filter SLAM
Graph-Based SLAM (Graph SLAM)	-	Графовый SLAM
ICP	-	Итеративный алгоритм ближайших точек (Iterative Closest Point)

ORB	- Комбинированный дескриптор (Oriented FAST and Rotated BRIEF)
RANSAC	- RANdom SAmple Consensus
SIFT	- Преобразование признаков, инвариантное к масштабу (Scale Invariant Feature Transform)
SLAM	- Одновременная навигация и построение карты (Simultaneous localization and mapping)
SURF	- Робастные ускоренные признаки (Speeded-Up Robust Features)
S-EKF SLAM	- Метод решения задачи одновременной навигации и картографирования на основе расширенного фильтра Калмана и семантических маркеров (Semantic Extended Kalman Filter SLAM)
YOLO	- Детектор объектов (You Only Look Once)
Visual SLAM	- Визуальный SLAM

Список использованных источников

1. Rui, T. Object tracking using particle filter in the wavelet subspace / T. Rui, Q. Zhang, Y. Zhou, J. Xing // *Neurocomputing*. – 2013. – Vol. 119. – P. 125–130.
2. Dissanayake, G. A Solution to the Simultaneous Localization and Map Building (SLAM) Problem / G. Dissanayake, P. Newman, S. Clark, H.-F. Durrant-Whyte, M. Csorba // *IEEE Transactions on Robotics and Automation*. – 2001. – Vol. 17 (3). – P. 229-241.
3. Zhou, S.K. Visual tracking and recognition using appearance-adaptive models in particle filters / S.K. Zhou, R. Chellappa, B. Moghaddam // *IEEE Transaction on Image Processing*. – 2004. – Vol. 13 (11). – P. 1491–1506.
4. Montemerlo, M. FastSLAM 2.0: An improved particle filtering algorithm for simultaneous localization and mapping that provably converges / M. Montemerlo, S. Thrun, D. Roller, B. Wegbreit // *Proceedings of the 18th International Joint Conference on Artificial Intelligence (Acapulco, Mexico August 2003)*. – 2003. – P. 1151-1156.
5. Estrada, C. Hierarchical SLAM: Real-time accurate mapping of large environments / C. Estrada, J. Neira, J.D. Tardós // *IEEE Transactions On Robotics*. – 2005. – Vol. 21 (4). – P. 588–596.
6. Se, S. Mobile robot localization and mapping with uncertainty using scaleinvariant visual landmarks / S. Se, D. Lowe, J. Little // *The international Journal of robotics Research*. – 2002. – Vol. 21 (8). – P.735–758.
7. Mur-Artal, R. Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras / R. Mur-Artal, J. D. Tardós // *IEEE Transactions on Robotics*. – 2017. – Vol. 33 (5). – P. 1255-1262.
8. Vokhmintcev, A.V. Reconstruction of three-dimensional map based on closed form solution of variational problem of multi-sensor data registration / A.V. Vokhmintcev, A.V. Melnikov, K.V. Mironov, V.V. Burlutskiy // *Doklady Mathematics*. – 2019. – Vol. 99 (1). – P. 1–5 (Impact Factor: 0.507, indexed in WoS, Scopus).

Вохминцев, А.В. Реконструкция трехмерных сцен на основе точных решений вариационной задачи регистрации мультисенсорных данных / А.В. Вохминцев, А.В. Мельников, К.В. Миронов, В.В. Бурлуцкий // Доклады академии наук. Математика. – 2019. – Т. 99 (1). – С. 1-5.

9. Kober, V. Recursive calculation of discrete cosine transform for low resolution signal processing / V. Kober, A. Vokhmintcev, A. Melnikov // Doklady Mathematics. – 2015. – Vol. 461 (3). – P. 257–260. (Impact Factor: 0.507, indexed in WoS, Scopus).

Кобер, В.И. Рекурсивное вычисление дискретного косинусного преобразования для обработки сигналов с пониженным разрешением / В.И. Кобер, А.В. Вохминцев, А.В. Мельников // Доклады академии наук. Математика. – 2015. – Т. 461 (3). – С. 257–260.

10. Vokhmintcev, A.V. Face recognition based on matching algorithm with recursive calculation of local oriented gradient histogram / A.V. Vokhmintcev, I.V. Sochenkov, V.V. Kuznetsov, D.V. Tikhonkikh // Doklady Mathematics. –2016. – Vol. 466 (3). – P. 453–459. (Impact Factor: 0.507, indexed in WoS, Scopus).

Вохминцев, А.В. Распознавание лиц на основе алгоритма сопоставления изображений с рекурсивным вычислением гистограмм направленных градиентов / А.В. Вохминцев, И.В. Соченков, В.В. Кузнецов, Д.В. Тихоньких // Доклады академии наук. Математика. – 2016. – Т. 466 (3). –С. 453–459.

11. 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 (Impact Factor: 0.21, indexed in WoS, Scopus).

12. Makovetskii, A. Frequency analysis of gradient descent method and accuracy of iterative image restoration/ Makovetskii, A. Vokhmintsev, V. Kober, V. Kuznetsov // Communications in Computer and Information Science. –2015. – Vol. 542. P. 109-118. (Impact Factor: 0.21, indexed in WoS, Scopus).

13. Вохминцев, А.В. Комбинированные методы навигации и составления карты на основе решения вариационной задачи точка-плоскость ICP 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. (Indexed in Scopus).

14. 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.

15. 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. (Indexed in Scopus).

16. 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.

17. 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. – №.2. – С. 16-26.

18. 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. (Indexed in Scopus).

19. 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. (Indexed in Scopus).

20. 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 (Indexed in WoS, Scopus).

21. 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 (Indexed in WoS, Scopus).

22. 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 (Indexed in WoS, Scopus).

23. 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 (Indexed in WoS, Scopus).

24. 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 (Indexed in WoS, Scopus).

25. 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 (Indexed in WoS, Scopus).

26. 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 (Indexed in WoS, Scopus).

27. 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 (Indexed in WoS, Scopus).

28. 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 А.В. Вохминцев // Труды института системного анализа РАН. – 2019. – Т.69 (4). – С. 3-12.

29. 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 (Indexed in WoS, Scopus).

30. 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 (Indexed in Scopus).

31. 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.

32. 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.

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 (Baden-Baden, 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. 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.

36. 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 (Indexed in WoS, Scopus).

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

38. 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.

39. 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.

40. 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.

41. 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.

42. 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.

43. 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.

44. Legacy Robots (Boston Dynamics). [Электронный ресурс]. Режим доступа: <https://www.bostondynamics.com/l3> (дата обращения: 01.12.2019).

45. Rollin' Justin (The Institute of Robotics and Mechatronics). [Электронный ресурс]. Режим доступа: <http://www.dlr.de/rmc/rm/en/desktopdefault.aspx/tabid-11427/#gallery/29202> (дата обращения: 01.12.2019).

46. Walking Robots (The National Centre of Competence in Research (NCCR) Robotics). [Электронный ресурс]. Режим доступа: <https://www.nccr-robotics.ch/research-areas/mobile-rescue-robots/walking-robots> (дата обращения: 01.12.2019).

47. Cobot Sawyer (Rethink Robotics GmbH). [Электронный ресурс]. Режим доступа: <http://www.rethinkrobotics.com> (дата обращения: 01.12.2019).
48. Nextage (Kawada Robotics). [Электронный ресурс]. Режим доступа: <http://nextage.kawada.jp/en/> (дата обращения: 01.12.2019).
49. Rocha, R. Cooperative multirobot systems: a study of vision-based 3-D mapping using information theory / R. Rocha, J. Dias, A. Carvalho A. // *International Journal of Robotics and Autonomous*. – 2005. – Vol . – P.282–311.
50. Hertzberg, C. Experiences in building a visual slam system from open source components / C. Hertzberg, R. Wagner, O. Birbach // *IEEE Proceedings of the International Conference on Robotics and Automation (Shanghai, China, May 2011)*. – 2011. – P. 2644-2651.
51. Montemerlo, M. FastSLAM: A factored solution to the simultaneous localization and mapping problem / M. Montemerlo, S. Thrun, D. Koller, B. Wegbreit // *Proceedings of the AAAI National Conference on Artificial Intelligence (Edmonton, Alberta, Canada, July - August 2002)*. – 2002. – P. 593–598.
52. Xiangzhang, C. Map Building Algorithm with High-precision Measurement Based on UFastSLAM / Xiangzhang C., Yongyin H., Yong W. // *Advances in Information Sciences and Service Sciences*. – 2012. – Vol. (13) P. 40-46.
53. Thrun, S. Probabilistic Robotics / S. Thrun, W. Burgard, D. Fox. – Cambridge: MIT Press. 2005. – 672 P.
54. Henry, P. RGB-D mapping: Using depth cameras for dense 3D modeling of indoor environments / P. Henry, M. Krainin, E. Herbst // *Springer Tracts in Advanced Robotics: Proceedings of the 12th International Symposium on Experimental Robotics (Delhi, India, December 2010)*. – 2014. – Vol. 79. – P. 477-491.
55. Endres, F. An evaluation of the rgb-d slam system / F. Endres, J. Hess, N., Engelhard, J. Sturm // *IEEE Proceedings of the International Conference on Robotics and Automation (Saint Paul, MN, USA, May 2012)*. – 2012. – P. 1- 6.
56. Fioraio, N. Realtime visual and point cloud SLAM / N. Fioraio, K. Konolige // *Proceedings of the 2nd RSS Workshop on RGB-D: Advanced reasoning with Depth Cameras at Robotics (Los Angeles, USA, June 2011)*. – 2011. – P. 477-491.

57. Thrun, S. Integrating Grid-Based and Topological Maps for Mobile Robot Navigation / S. Thrun, A. Bucken // Proceedings of the 13th National Conference on Artificial Intelligence, AAAI (Portland, Oregon, August 1996) 1996, P. 944-950.
58. Newman, P. Navigating, recognising and describing urban spaces with vision and laser / P. Newman, G. Sibley, M. Smith, M. Cummins, A. Harrison, C. Mei // The International Journal of Robotics Research. – 2009. – Vol .28. –P. 1406–1433.
59. Nister, D. Visual Odometry / D. Nister, O. Naroditsky, J. Bergen // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR (Washington, DC, USA, June -July 2004). – 2004. – P. 477-491.
60. Nister, D. Visual Odometry for Ground Vehicle Applications / D. Nister, O. Naroditsky, J. Bergen // Journal of Field Robotics. – 2006. – Vol. 23(1). –P.274-283.
61. Nister, D., Scalable Recognition with a Vocabulary Tree / D. Nister, H. Stewenius // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR (New York, NY, USA, June 2006). – 2006.– Vol.2. – P. 2161-2168.
62. Pollefeys, M. Detailed Real-Time Urban 3D Reconstruction from Video / M. Pollefeys, D. Nister, J. M. Frahm, A. Akbarzadeh, P. Mordohai, B. Clipp, C. Engels, D. Gallup, S. J. Kim, P. Merrell, C. Salmi, S. Sinha, B. Talton, L. Wang L., Q. Yang, H. Stewenius, R. Yang, G. Welch, H. Towles // International Journal of Computer Vision (IJCV). – 2008. – Vol.78 (2-3). – P. 143–167.
63. Akbarzadeh, A. Towards urban 3D reconstruction from video / A. Akbarzadeh, J.-M., Frahm, P. Mordohai, B. Clipp, C. Engels, D. Gallup [et al] // Proceedings of the 3rd International Symposium on 3D Data Processing, Visualization and Transmission, 3DPVT (Chapel Hill, NC, USA, June 2006). – 2006. – P.1–8.
64. Merrell, M. Real-Time Visibility-Based Fusion of Depth Maps / M. Merrell, A. Akbarzadeh, L. Wang, P. Mordohai, J.-M. Frahm, R. Yang, D. Nister, M. Pollefeys // Proceedings of the 11th International Conference on Computer Vision, ICCV (Rio de Janeiro, Brazil, October 2007). – 2007. – P.1–8.
65. Konolige, K. Large-scale map making / K. Konolige K. // Proceedings of the 19th National Conference on Artificial Intelligence, AAAI (San Jose, California, July 2004). – 2004. – P.457–463.

66. Konolige K., Projected texture stereo / K. Konolige // Proceedings of the IEEE International Conference on Robotics and Automation, ICRA (Anchorage, AK, USA, May 2010). – 2010. – P. 148–155.
67. Clemente, L. Mapping large loops with a single hand-held camera / L. Clemente, A. Davison, I. Reid, J. Neira, J. Tardos // Proceedings of Robotics: Science and Systems III, RSS (Atlanta, Georgia, USA, June 2007). – 2007. – Vol.2. – P. 1–8.
68. Snavely, N. Photo tourism: Exploring photo collections in 3D / N. Snavely, S. Seitz, R. Szeliski // ACM Transactions on Graphics. – 2006. – Vol .25 (3). – P. 835–846.
69. Furukawa, Y. Accurate, Dense, and Robust Multiview Stereopsis / Y. Furukawa, J. Ponce // IEEE Transactions Pattern Analysis and Machine Intelligence. – 2010. – Vol.32 (8). – P.1362–1376.
70. Furukawa, Y. Reconstructing building interiors from images / Y. Furukawa, B. Curless, S. Seitz, R. Szeliski // Proceedings of the IEEE 12th International Conference on Computer Vision (ICCV (Kyoto, Japan, September – October 2009). – 2010. – P. 80–87.
71. Besl, P. A method for registration of 3-D shapes / P. Besl, N. McKay // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1992. – Vol.14 (2). – P.239–256.
72. Thrun, S. A real-time algorithm for mobile robot mapping with applications to multi-robot and 3D mapping / S. Thrun, W. Burgard, D. Fox // Proceedings of the IEEE International Conference on Robotics and Automation, ICRA (San Francisco, CA, USA, April 2000). – 2000. – P.321–328.
73. Rusinkiewicz, S. Efficient variants of the ICP algorithm / S. Rusinkiewicz, M. Levoy // Proceedings of the IEEE International Conference on 3-D Digital Imaging and Modeling, 3DIM (Quebec City, Canada, May-June 2001). – 2002. – P.145–152.
74. Chen, Y. Object modeling by registration of multiple range images / Y. Chen, G. Medioni // Image and Vision Computing. – 1992. – Vol.10. – P.145–155.

75. Segal, A. Generalized-ICP / A. Segal, D. Haehnel, S. Thrun // Proceedings of Robotics: Science and Systems, RSS (Seattle, USA, June - July 2009). – 2009, Vol.2 (4). – P. 1-8.
76. Triggs, B. Bundle adjustment—a modern synthesis / B. Triggs, P. McLauchlan, R. Hartley, A. Fitzgibbon // Vision Algorithms: Theory and Practice. – 2000. – Vol.1883. – P. 153–177.
77. Lourakis, M. SBA: A software package for generic sparse bundle adjustment / M. Lourakis, A. Argyros // ACM Transactions on Mathematical Software. – 2009. – Vol. 36. – P.1–30.
78. Konolige, K. Sparse sparse bundle adjustment / K. Konolige // Proceedings of the British Machine Vision Conference, BMVC (Aberystwyth, UK, August-September 2010). – 2010. – P. 102.1-102.11.
79. Konolige, K. FrameSLAM: From bundle adjustment to real-time visual mapping / K. Konolige, M. Agrawal // IEEE Transactions on Robotics. – 2008, Vol. 25 (5). – P.1066–1077.
80. Magnusson, M. Registration for Autonomous Mining Vehicles Using 3DNDT / M. Magnusson, A. Lilienthal, T. DuckettScan // Journal of Field Robotics. – 2007. – Vol. 24(10). – P. 803–827.
81. Stentz, A. Optimal and Efficient Path Planning for Partially-Known Environments / A. Stentz // Proceedings of the IEEE International Conference on Robotics and Automation (San Diego, CA, USA, May 1994). – 2002. – P. 3310–3317.
82. Stentz, A. The Focussed D* Algorithm for Real-Time Replanning / A. Stentz // Proceedings of the International Joint Conference on Artificial Intelligence (Montreal, Canada August 1995). – 1995. – P. 1652–1659.
83. Ulaş, C. 3D Multi-Layered Normal Distribution Transform for Fast and Long Range Scan Matching / C. Ulaş, H. Temeltaş // Journal of Intelligent & Robotic Systems. – 2013. Vol. 71(1). – P. 85–108.
84. Edelsbrunner, H. Algorithms in Combinatorial Geometry / H. Edelsbrunner EATCS Monographs on Theoretical Computer Science. – Berlin/Heidelberg: Springer-Verlag. 1987. – Vol. 10. – 423 P.

85. Wooden, D.T. Graph-based Path Planning for Mobile Robots: PhD thesis / D.T. Wooden // Georgia Institute of Technology. – Atlanta, USA. – 2006. – P. 114.
86. Branicky, M. Quasi randomized path planning / M. Branicky, S. Lavalle, K. Olson, L. Yang // Proceedings of the IEEE International Conference on Robotics and Automation (Seoul, South Korea, May 2001). – 2003. – Vol. 2. – P. 1481-1487.
87. M.S. Branicky Sampling-based planning, control, and verification of hybrid systems / M.S. Branicky, M.M. Curtiss, J. Levine, S. Morgan // IFAC Proceedings of the 16th IFAC World Congress (Prague, Czech, July 2005). – 2005. – Vol. 38 (1). – Pages 271-276.
88. Yap, P. Grid-based path-finding / P. Yap // Proceedings of the 15th Conference of the Canadian Society for Computational Studies of Intelligence on Advances in Artificial Intelligence (Calgary, Canada, May 2002). – 2002. – P. 44-55.
89. Grisetti, G. A tutorial on graph-based SLAM / G. Grisetti, R. Kummerle, C. Stachniss, W. Burgard // IEEE Intelligent Transportation Systems Magazine. – 2010 – Vol. 2(4). – P. 31-43.
90. Newcombe, R. A. KinectFusion: Real-time dense surface mapping and tracking / R.A. Newcombe, S. Izadi, O. Hilliges, D. Molyneaux, D. Kim, A.J. Davison, P. Kohli, J. Shotton, S. Hodges, A. Fitzgibbon // Proceedings IEEE 10th International Symposium on Mixed and Augmented Reality, ISMAR (Basel, Switzerland, October 2011). – 2012. – P. 127–136.
91. Whelan, T. Kintinuous: Spatially extended kinectfusion / T. Whelan, M. Kaess, M. Fallon, H. Johannsson, J. Leonard, J. McDonald // RSS Workshop on RGB-D: Advanced Reasoning with Depth Cameras (Pittsburgh, USA, July 2012). – 2012. P. 1-8. https://www.ri.cmu.edu/pub_files/2012/7/Whelan12rssw.pdf
92. Sturm, J. A benchmark for the evaluation of RGB-D SLAM systems / J. Sturm, N. Engelhard, F. Endres, W. Burgard, D. Cremers // Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (Vilamoura, Portugal, October 2012). – 2012. – P.573–580.

93. Henry, P. RGB-D mapping: Using depth cameras for dense 3D modeling of indoor environments / P. Henry, M. Krainin, E. Herbst, X. Ren, D. Fox // *The International Journal of Robotics Research*. 2012 – Vol. 31(5). – P. 647-663.
94. Fuentes-Pacheco, J. Visual simultaneous localization and mapping: a survey / J. Fuentes-Pacheco, J. Ruiz-Ascencio, J.M. Rendón-Mancha // *Artificial Intelligence Review*. 2015 – Vol. 43(1). – P. 55-81.
95. Lowe, D. G. Distinctive image features from scale-invariant keypoints / D.G. Lowe // *International journal of computer vision*. – 2004. – Vol. 60(2). – P. 91-110.
96. Bay, H., Speeded-up robust features (SURF) / H. Bay, A. Ess, T. Tuytelaars, L. Van Gool // *Computer Vision and Image Understanding*. – 2008 – Vol. 110 (3). P. 346-359.
97. Rublee, E. ORB: An efficient alternative to SIFT or SURF / E. Rublee, V. Rabaud, K. Konolige, G. Bradski // *Proceedings of the IEEE International Conference on Computer Vision (Barcelona, Spain, November 2011)*. – 2012. – P. 2564-2571.
98. Raguram, R. A comparative analysis of RANSAC techniques leading to adaptive real-time random sample consensus / R. Raguram, J. Frahm, M. Pollefeys // *Lecture Notes in Computer Science: Proceedings of the European Conference on Computer Vision (Marseille, France, October 2008)*. – 2008. – Vol. 5303. –P. 500–513.
99. Thompson, E.H. On exact linear solution of the problem of absolute orientation / E.H. Thompson // *Photogrammetria*. – 1958. – Vol. 13. –P. 163–178.
100. Horn, B. Closed-Form Solution of Absolute Orientation Using Unit Quaternions / B. Horn // *Journal of the Optical Society of America. A*. – 1987. – Vol. 4(4). – P. 629–642.
101. Horn, B. Closed-form Solution of Absolute Orientation Using Orthonormal Matrices / B. Horn, H. Hilden, S. Negahdaripour // *Journal of the Optical Society of America A*. – 1988. – Vol. 5(7). – P. 1127-1135.
102. Du, S. An extension of the ICP algorithm considering scale factor / S. Du, N. Zheng, S. Ying, Q. You, Y. Wu // *Proceedings of the 14th IEEE International Conference on Image Processing, ICIP (San Antonio, USA, September – October 2007)*. – 2007. – P.193–196.

103. Du, S. Affine Registration of Point Sets Using ICP and ICA / S. Du, N. Zheng, G. Meng, Z. Yuan // IEEE Signal Processing Letters. – 2008. – Vol. 15. – P. 689-692.
104. Du, S. Affine iterative closest point algorithm for point set registration / S. Du, N. Zheng, S. Ying, J. Liu // Pattern Recognition Letters. – 2010. – Vol. 31. –P. 791–799.
105. Tihonkih, D. The iterative closest points algorithm and affine transformations / D. Tihonkih, A. Makovetskii, V. Kuznetsov // CEUR Workshop Proceedings of the 4th International Conference of Analysis of Images, Social Networks, and Texts, AIST 2016 (Yekaterinburg, Russia, April 2016). – 201. – P. 349-356.
106. Linear least-squares optimization for point-to-plane ICP surface registration [Текст]: Technical Report: TR04-004 / Department of Computer Science, University of North Carolina; K. L. Low. – Chapel Hill, 2004. – 3 P.
107. Van der Sande, C. Assessment of relative accuracy of AHN-2 laser scanning data using planar features / C. Van der Sande, S. Soudarissanane, K. Khoshelham // Sensors. – 2010. – Vol. 10. – P. 8198– 8214.
108. Curless, B. A volumetric method for building complex models from range images / B. Curless, M. Levoy // Proceedings of the 23rd Annual Conference on Computer graphics and interactive techniques (New Orleans, USA, August 1996). – 1996. – P. 3030-312.
109. Wurm, K.M. OctoMap: A probabilistic, flexible, and compact 3D map representation for robotic systems / K.M. Wurm, A. Hornung, M. Bennewitz, C. Stachniss, W. Burgard // Proceedings of the ICRA Workshop on Best Practice in 3D Perception and Modeling for Mobile Manipulation (Anchorage, USA, May 2010). – 2010. – P. 1-8.
110. Weise, T. Inhand scanning with online loop closure / T. Weise, T. Wismer, B. Leibe, L. Gool L. // Proceedings of the IEEE 12th International Workshop on 3D Digital Imaging and Modeling, 3DIM (Kyoto, Japan, September – October 2009). – 2010. –P. 1630 -1637.

111. Hernandez, C. Probabilistic visibility for multi-view stereo / C. Hernandez, G. Vogiatzis, R. Cipolla // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR (Minneapolis, USA, June 2007). – 2007. – P. 1-8.
112. Clarke, R. Human identification in information systems: Management challenges and public policy issues / R. Clarke // Information Technology & People. – 1994. – Vol. 7(4). – P. 6-37.
113. Jain, A.K. Biometrics: Personal Identification in Networked Society / A.K. Jain, R.M. Bolle, M. Ruud, S. Pankanti. – Berlin/Heidelberg: Springer-Verlag. 2006. – 411 P.
114. ГОСТ Р ИСО/МЭК 19794-1-2008. Автоматическая идентификация. Идентификация биометрическая. Форматы обмена биометрическими данными. [Текст]. – Структура. 2009-01-01. – М.: Стандартинформ: Изд-во стандартов, 2009. – V, 42 с.: ил.; 18.
115. ГОСТ Р ИСО/МЭК 19794-5-2006. Автоматическая идентификация. Идентификация биометрическая. Форматы обмена биометрическими данными. [Текст]. – Данные изображения лица. 2007-01-01. – М.: Стандартинформ: Изд-во стандартов, 2006. – I, 20 с.: ил.; 7.
116. ГОСТ Р ИСО/МЭК 19795-2. Автоматическая идентификация. Идентификация биометрическая. Эксплуатационные испытания и протоколы испытаний в биометрии. [Текст]. – Методология проведения технологического и сценарного испытаний. 2009-01-01. – М.: Стандартинформ: Изд-во стандартов, 2009. – II, 47 с.: ил.; 2.
117. Болл, Р. М. Руководство по биометрии / Р.М. Болл, Д.Х. Коннел, Ш. Шарат, К. Ратхт Налини, У. Эндрю. – М: Техносфера. 2007. – 368 С.
118. Blanz, V. Face Recognition Based on Frontal Views Generated from Non-frontal Images / V. Blanz, P. Grother, J. Phillips, T. Vetter // Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR (San Diego, USA, June 2005). – 2005. – P. 454–461.
119. Belhumeur, R.H. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection / R.H. Belhumeur, J.R Hespanha, D.J. Kriegman // IEEE

Transactions on Pattern Analysis and Machine Intelligence. – 1997. – Vol. 19 (7). – P.711-720.

120. Senior, A.W. Recognizing faces in broadcast video / A.W. Senior // Proceedings of the International Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems. In Conjunction with ICCV'99 (Corfu, Greece, September 1999). – 2002. – P. 105-110.

121. Kim, T. Locally Linear Discriminant Analysis for Multimodally Distributed Classes for Face Recognition / T. Kim T., J. Kittler // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2005. – Vol. 27 (3). – P. 318–327.

122. Khaleghian S. Face recognition across large pose variations via Boosted Tied Factor Analysis / S. Khaleghian, H.R. Rabiee, M.H. Rohban // Proceedings of the IEEE Workshop on Applications of Computer Vision, WACV (Kona, USA, January 2011). –2011. –P. 190 – 195.

123. Заварикин, Д. Принципы построения и преимущества системы 3D-распознавания лиц / Д. Заварикин // Технологии защиты. – 2010. – №4. –С. 72-76.

124. Papatheodorou, T. Evaluation of 3D face recognition using registration and PCA / T. Papatheodorou, D. Rueckert // Proceedings of the 5th International Conference on Audio- and Video-Based Biometric Person Authentication (Hilton Rye Town, NY, USA, July 2005). – 2005. – P. 997–1009.

125. Ojala, T. A Comparative Study of Texture Measures with Classification Based on Feature Distributions / T. Ojala, M. Pietikainen, D. Harwood // Pattern Recognition. – 1996. – Vol. 29. – P. 51-59.

126. Moreno, A. Face recognition using 3D surface extracted descriptors / A. Moreno, A. Sanchez, J. Velez, F. Diaz // Proceedings of the In Irish Machine Vision and Image Processing Conference (Northern Ireland September 2003). – 2003. – P. 1-8.

127. Lee, Y. Local feature based 3D face recognition / Y. Lee, H. Song, U. Yang, H. Shin K. Sohn // Proceedings of the 5th International Conference : Audio- and Video-Based Biometric Person Authentication, AVBPA (Hilton Rye Town, NY, USA, July 2005). – 2005. P. – 909-918.

128. Bronstein, A. Three-dimensional face recognition / A. Bronstein, M. Bronstein, R. Kimmel // *International Journal of Computer Vision*. – 2005. – Vol. 64. – P. 5–30.
129. Горин, С.В. Использование информации о динамике изменений человеческого лица для решения задач распознавания и классификации / С.В. Горин, А.С. Сулимов // *Инженерный журнал: наука и инновации*. – 2013. – Т. 6 (18). – С. 1-15.
130. Gu, J. Action and gait recognition from recovered 3-D human joints / J. Gu, X. Ding, S. Wang, Y. Wu // *IEEE Trans. Sys. Man Cyber. Part B*. – 2010. – Vol. 40. – P. 1021–1033.
131. Medioni, G. Identifying Non-cooperative Subjects at a Distance Using Face Images and Inferred Three-Dimensional Face Models / G. Medioni, J. Choi, C.-H. Kuo, D. Fidaleo // *IEEE Trans. Sys. Man Cyber. Part A*. – 2009. – Vol. 39 (1). – P. 12-24.
132. Chen, S.Y. Kalman filter for robot vision: a survey / S. Y. Chen // *IEEE Transactions on Industrial Electronics* – 2012. – Vol. 59. – P. 4409–4420.
133. Yilmaz, A. Object tracking: A survey / A. Yilmaz, O. Javed, M. Shah // *ACM Computer Survey*. – 2006. – Vol. 38 (4). – P. 1– 45.
134. Doucet, A. On sequential Monte Carlo sampling methods for Bayesian filtering / A. Doucet, S. Godsill, C. Andrieu // *Statistics and Computing*. – 2010. – Vol. 10. – P. 197–208.
135. Zhou, S.K. Visual tracking and recognition using appearanceadaptive models in particle filters / S. K. Zhou, R. Chellappa, B. Moghaddam // *IEEE Transaction on Image Processing*. – 2004. – Vol. 13 (11). – P. 1491– 1506.
136. Rui, T. Object tracking using particle filter in the wavelet subspace / T. Rui, Q. Zhang, Y. Zhou, J. Xing // *Neurocomputing*. – 2013. – Vol. 119. – P. 125–130.
137. Isard, M. Condensation - conditional density propagation for visual tracking / M. Isard, A. Blake // *International Journal of Computer Vision*. – 1998. – Vol. 29. – P. 5–28.
138. Reid, D. An algorithm for tracking multiple targets / D. Reid // *IEEE Transaction on Automatic Control*. – 1979. – Vol. 24 (6). – P. 843–854.

139. Cox, I. An efficient implementation of Reid's multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking / I. Cox, S. Hingorani // IEEE Transaction on Pattern Analysis and Machine Intelligence. – 1996. – Vol.18 (2). – P. 138–150.
140. Comaniciu, D. Kernel-based object tracking / D. Comaniciu, V. Ramesh, P. Meer // IEEE Transaction on Pattern Analysis and Machine Intelligence. – 2003. – Vol. 25 (5). – P. 564–577.
141. Jepson, A. Robust online appearance models for visual tracking / A. Jepson, D.Fleet, T. ElMaraghi // IEEE Transaction on Pattern Analysis and Machine Intelligence. – 2003. – Vol. 25 (10). – P. 1296–1311.
142. Lucas, B. An iterative image registration technique with an application to stereo vision / B. Lucas, T. Kanade // Proceedings of the 7th International joint conference on Artificial intelligence (Vancouver, Canada, August 1981). – 1981. – Vol. 2. – P. 674–679, 1981.
143. Mozerov, M. Constrained optical flow estimation as a matching problem / M. Mozerov // IEEE Transactions on Image Processing. – 2013. – Vol. 22 (5). – P. 2044–2055.
144. Xu, L. Motion detail preserving optical flow estimation / L. Xu, J. Jia, Y. Matsushita // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2012. – Vol. 34 (9). – P.1744 -1757.
145. Comaniciu, D. Real-time tracking of non-rigid objects using mean shift / D. Comaniciu, V. Ramesh, P. Meer // Proceedings IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2000 (Hilton Head Island, SC, USA, June 2000). – 2002. – P.142–151.
146. Leichter, I. Mean shift tracking with multiple reference color histograms / I. Leichter, M. Lindenbaum, E. Rivlin // Computer Vision and Image Understanding. – 2010. – Vol. 114. – P. 400–408.
147. Zhou, H. Object tracking using SIFT features and mean shift / H. Zhou, Y. Tuan, C. Shi // Computer Vision and Image Understanding. – 2009. – Vol. 113. – P. 345–352.

148. Lowe, D.G. Object recognition from local scale-invariant features / D.G. Lowe // Proceedings of the International Conference on Computer Vision (Kerkyra, Corfu, Greece, September 1999). – 1999. – Vol. 2. – P. 1150–1157.
149. Kalal, Z. Tracking-learning-detection / Z. Kalal, K. Mikolajczyk, J. Matas // IEEE Transaction on Pattern Analysis and Machine Intelligence. – 2012. – Vol. 34 (7). – P. 1409–1422.
150. Manzurov, T. Optical correlator based target detection, recognition, classification, and tracking / T. Manzurov, J. Zeller, S. Serati // Applied Optics. 2012. – Vol. 51. – P. 4976–4983.
151. Rong Li, X. Survey of maneuvering target tracking. Part I. dynamic models / X. Rong Li, V. P. Jilkov // IEEE Transactions on Aerospace and Electronics Systems. – 2003. – Vol. 39 (4). P. 1333–1364.
152. Diaz-Ramirez, V.H. Multiclass pattern recognition using adaptive correlation filters with complex constraints / V.H. Diaz-Ramirez, O.G. Campos-Trujillo, V. Kober // Optical Engineering. – 2012. – Vol. 51 (3). – P. 037203-12.
153. Kober, V. Discrimination of similar objects using partial phase information / V. Kober, V. Karnaukhov, O. Milukova // Proceedings of the 11th International Conference on Pattern Recognition and Image Analysis: New Information Technologies, PRIA (Samara, Russia, September 2013). –2013. – P. 217-220.
154. Recognition of Noncooperative Individuals Using 3-D Face Biometric (University of Southern California). [Электронный ресурс]. Режим доступа: <http://www-bcf.usc.edu/~nevatia/> (дата обращения: 01.12.2019).
155. Multimodal Biometric Fusion With Predictive Quality Metrics (General Electric Company). [Электронный ресурс]. Режим доступа: <https://www.innovatrics.com/contact/> (дата обращения: 01.12.2019).
156. Sensors, Surveillance and Biometrics Center of Excellence (Department of Defense Research and Development Laboratory, Massachusetts Institute of Technology (MIT) Lincoln Laboratory). [Электронный ресурс]. Режим доступа: <https://www.ll.mit.edu/> (дата обращения: 01.12.2019).

157. Image Stabilized Binoculars With Integrated 3-D Facial Recognition Imaging Capabilities (StereoVision Imaging). [Электронный ресурс]. Режим доступа: <http://stereovisioninc.com/> (дата обращения: 01.12.2019).

158. Face Recognition From Video (University of Notre Dame). [Электронный ресурс]. Режим доступа: <https://cvrl.nd.edu/projects/data/> (дата обращения: 01.12.2019).

159. Facial and Iris-Based Biometric System (New Jersey Institute of Technology). [Электронный ресурс]. Режим доступа: <https://frvp.njit.edu/index.html>; <https://centers.njit.edu/research-areas/data-science-and-information-technology/> (дата обращения: 01.12.2019).

160. Efficient, Field-Optimized, Multimodal Biometric System (International Biometric Group). [Электронный ресурс]. Режим доступа: <https://biometricgroup.com> (дата обращения: 05.11.2018).

161. Facenorm: Normalization Plug-In for Improved Face Recognition of Noncooperative Individuals (Carnegie Mellon University). [Электронный ресурс]. Режим доступа: <https://www.cylab.cmu.edu/> (дата обращения: 01.12.2019).

162. Вокорд. Системы безопасности и видеонаблюдения; https://vocord.ru/directions/face_detection/ (дата обращения: 01.12.2019).

163. Stanley (DARPA, Stanford Artificial Intelligence Lab, the Stanford Racing Team, Volkswagen Electronics Research Laboratory). [Электронный ресурс]. Режим доступа: <http://ai.stanford.edu/#faculty>; <https://cs.stanford.edu/group/roadrunner//old/index.html> (дата обращения: 01.12.2019).

164. Robots Podcast: 3D SLAM (Department of Defense Research and Development Laboratory, Massachusetts Institute of Technology (MIT) Lincoln Laboratory). [Электронный ресурс]. Режим доступа: <http://groups.csail.mit.edu/rrg/>; <http://groups.csail.mit.edu/locomotion/>; <http://interactive.mit.edu/> (дата обращения: 01.12.2019).

165. Autonomous and Perceptive Systems research (University of Groningen). [Электронный ресурс]. Режим доступа: <http://www.teamborg.nl/> (дата обращения: 01.12.2019).

166. Active 3-D Visual SLAM (General Electric Company). [Электронный ресурс]. Режим доступа: <https://www.ge.com/research/technology-domains/robotics-autonomous-systems> (дата обращения: 01.12.2019).

167. SLAM using vision (Imperial College London). [Электронный ресурс]. Режим доступа: <https://www.imperial.ac.uk/dyson-robotics-lab/projects/codeslam/> (дата обращения: 01.12.2019).

168. WPI Robotics Engineering C Squad –WRECS (Worcester Polytechnic Institute). [Электронный ресурс]. Режим доступа: <https://www.wpi.edu/academics/departments/robotics-engineering> (дата обращения: 01.12.2019).

169. JPL / UCSB / Caltech (Jet Propulsion Laboratory, Pasadena). [Электронный ресурс]. Режим доступа: <https://www.jpl.nasa.gov/> (дата обращения: 01.12.2019).

170. Team TRAC Labs (TRAC Labs Inc). [Электронный ресурс]. Режим доступа: <https://traclabs.com/> (дата обращения: 01.12.2019).

171. Volkswagen of America (Electronics Research Laboratory). [Электронный ресурс]. Режим доступа: <http://www.vwerl.com/our-work> (дата обращения: 01.12.2019).

172. Team IHMC (Institute for Human and Machine Cognition). [Электронный ресурс]. Режим доступа: <https://www.ihmc.us/research/biologically-inspired-robots> (дата обращения: 01.12.2019).

173. TORC (TORC, TU Darmstadt, Virginia Tech). [Электронный ресурс]. Режим доступа: <https://torc.ai/team-vigir-team/> (дата обращения: 01.12.2019).

174. Miramontes-Jaramillo, D. A novel Image Matching Algorithm Based on Sliding Histograms of Oriented Gradients / D. Miramontes-Jaramillo, V. Kober, V.H. Diaz-Ramirez, V. Karnaukhov // Journal of Communications Technology and Electronics. – 2014. – Vol. 59 (12). – P. 1446–1450.

175. Miramontes-Jaramillo, D. Multiple objects tracking with HOGs matching in circular windows / D. Miramontes-Jaramillo, V. Kober, V.H. Diaz-Ramirez //

Proceedings of the SPIE XXXVII Annual Meeting: Applications of Digital Image Processing (San Diego, USA, September 2014). – 2014. – Vol. 9217. – P. 92171N-8.

176. Echeagaray-Patron, B.A. A Method of face recognition using 3D facial surfaces / B.A. Echeagaray-Patron, V.I. Kober, V.N. Karnaukhov, V.V. Kuznetsov // Journal of Communications Technology and Electronics. – 2017. – Vol. 62(6). – P. 648-652.

177. Liu, B. New fast algorithms for the estimation of block motion vectors / B. Liu, A. Zaccarin // IEEE Transactions on Circuits and Systems for Video Technology. – 1993. – Vol. 3 (2). – P. 148–157.

178. Petitot, J. The neurogeometry of pinwheels as a sub-Riemannian contact structure / J. Petitot // Journal Physiology–Paris. – 2003. – Vol. 97 (2-3). – P. 265–309.

179. NYU Depth Dataset V2 (NYU CS department, Courant Institute of Mathematical Sciences). [Электронный ресурс]. Режим доступа: http://cs.nyu.edu/~silberman/datasets/nyu_depth_v2.html#raw_parts (дата обращения: 01.12.2019).

180. Yakovlev, K. HGA*, an efficient algorithm for path planning in a plane / K. Yakovlev // Scientific and Technical Information Processing. – 2010. – Vol. 37(6). – P. 438-447.

181. Bresenham, J. E. Algorithm for computer control of a digital plotter / J. E. Bresenham // IBM Systems Journal. – 1965. – Vol. 4 (1). – P.

182. ASL Datasets Repository (Autonomous Systems Lab, Swiss Federal Institute of Technology Zurich). [Электронный ресурс]. Режим доступа: <https://projects.asl.ethz.ch/datasets/> (дата обращения: 01.12.2019).

183. Redmon, J. YOLO9000: Better, Faster, Stronger / J. Redmon, A. Farhadi // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), (Las Vegas, USA, June 2016). – 2016. – P. 7263-7271.

184. Redmon, J. Yolov3: An incremental improvement / J. Redmon, A. Farhadi. – arXiv preprint arXiv:1804.02767. – 2018.

185. Wolf, D. Enhancing semantic segmentation for robotics: the power of 3-d entangled forests / D. Wolf, J. Prankl, M. Vincze // IEEE Robotics and Automation Letters. 2016 – Vol. 1(1). – P. 49–56.
186. Papon, J. Voxel cloud connectivity segmentation-supervoxels for point clouds / J. Papon, A. Abramov, M. Schoeler, F. Worgotter // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), (Portland, USA, June 2013). – 2013. – P. 2027–2034.
187. Antonello, M. Multi-view 3d entangled forest for semantic segmentation and mapping / M. Antonello, D. Wolf, J. Prankl, S. Ghidoni, E. Menegatti, M. Vincze // Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), (Brisbane, Australia, May 2018). – 2018 – P. 1855–1862.
188. Rother, C. Grabcut: Interactive foreground extraction using iterated graph cuts / C. Rother, V. Kolmogorov, A. Blake // ACM Transactions on Graphics (TOG). 2004 – Vol. 23(3). – P. 309–314.
189. Aguilar-Gonzalez, P.M. Design of correlation filters for pattern recognition using a noisy reference / P.M. Aguilar-Gonzalez, V. Kober // Optics Communications. 2012. – Vol. 285. – P. 574–583.
190. Diaz-Ramirez, V.H. Multiclass pattern recognition using adaptive correlation filters with complex constraints / V.H. Diaz-Ramirez, O.G. Campos-Trujillo, V. Kober // Optical Engineering. – 2012. – Vol. 51 (3). – P. 037203-12.
191. Pratt, W.K. Digital Image Processing / W.K. Pratt. – NY: Wiley & Sons. 2007.
192. Caulfield, H.D. Improved discrimination in optical character recognition / H.D. Caulfield, W.T. Maloney // Applied Optics. – 1969. – Vol. 8. – P. 2354-2356.
193. Chan, T. Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic Methods / T. Chan, J. Shen. – Philadelphia: Society for Industrial and Applied Mathematics, 2005. – 184 P.
194. Rosen, J.G. The gradient projection method for nonlinear programming. Part II, nonlinear constraints / J.G. Rosen // Journal of the Society for Industrial and Applied Mathematics. – 1961. – Vol. 9 (4). – P. 514-532.

195. Goldstein, T. The split Bregman method for L1 regularized problems / T. Goldstein, S. Osher // *SIAM Journal on Imaging Sciences*. – 2009. – Vol. 2 (2). – P. 323-343.
196. Cai, J. Split Bregman methods and frame based image restoration / J. Cai, S. Osher, Z. Shen // *Multiscale Modeling and Simulation*. – 2009. – Vol. 8 (2). –P. 337-369.
197. Osher, S. An iterative regularization method for total variation based image restoration / S. Osher, M. Burger, D. Goldfarb, J. Xu J., W. Yin // *Multiscale Modelling and Simulation*. – 2005. – Vol. 4. – P. 460-489.
198. Lopez-Martinez, L. Métodos de restauración de imágenes utilizando microscaneo / L. Lopez-Martinez, V. Kober. – Saarbrücken: AV Akademikerverlag. 2013. – P. 110.
199. Lopez-Martinez, J. L. Blind adaptive method for image restoration using microscanning / J.L. Lopez-Martinez, V. Kober // *IEICE Transactions on Information and Systems*. – 2012. – Vol. E95-D (3). – P. 280-284.
200. Lopez-Martinez, J.L Blind Image Restoration for a Microscanning Imaging System / J.L. Lopez-Martinez, V. Kober // *Image Restoration – Recent Advances and Applications*, ed. A. Histace. Chapter 13 in book. Rijeka: InTech Publishing, 2012 – P. 297-314. ISBN 978-953-51-0388-2.
201. López-Martínez, J.L. Accuracy of image restoration using microscanning imaging system / J.L. López-Martínez, V. Kober // *Proceedings of the SPIE XXXI Annual Meeting: Applications of Digital Image Processing (San-Diego, USA, August 2011)*. – 2011. – Vol. 8135. – P. 81351F-8.
202. López-Martínez, J.L. Image restoration of nonuniformly illuminated images with camera microscanning / J.L. López-Martínez, V. Kober // *Proceedings of the SPIE XXXIII Annual Meeting: Applications of Digital Image Processing (San-Diego, USA, September 2010)*. – 2010. – Vol. 7798. – P. 77982D-6, 2010.
203. Kober, V. Discrimination of similar objects using partial phase information / V. Kober, V. Karnaukhov, O. Milukova // *Proceedings of the. 11th International Conference on Pattern Recognition and Image Analysis (Minsk, Belarus)*. – 2013. – P. 117-120.

204. Picos, K. Object tracking under nonuniform illumination with adaptive correlation filtering / K. Picos, V.H. Díaz-Ramírez, V. Kober // Proceedings of the SPIE XXXVI Annual Meeting: Applications of Digital Image Processing (San Diego, USA, August 2013). – 2013. – Vol. 8856. – P. 88550H.
205. Kober, V. Image restoration with sliding sinusoidal transforms / V. Kober, I.A. Ovseevich // Pattern Recognition and Image Analysis. – 2008. – Vol. 18 (4). – P. 650-654.
206. Kober, V. Fast algorithms for the computation of sliding discrete Hartley transforms / V. Kober // IEEE Transactions on Signal Processing. – 2007. – Vol. 55 (6). – P. 2937 – 2944.
207. Kober, V. Fast algorithms for the computation of sliding sinusoidal transforms / V. Kober // IEEE Transactions on Signal Processing. – 2004. – Vol. 52 (6). – P. 1704-1710.
208. Jain, A.K. A sinusoidal family of unitary transforms / A.K Jain // IEEE Trans. Pattern Anal. Machine Intell. – 1979. – Vol. PAMI-1 (4). – P. 356–365.
209. Oppenheim, A.V. Discrete Time Signal Processing / A. V. Oppenheim, R. W. Shafer. – NJ: Prentice Hall, Englewood Cliffs; 1989. – 467 P.
210. Vitkus, R. Y. Recursive algorithms for local adaptive linear filtration / R. Y. Vitkus, L. P. Yaroslavsky // in Mathematical Research. – Berlin: Academy Verlag. –1987. – P.34-39.
211. Yip, P. On the shift property of DCT's and DST's / P. Yip, K. P. Rao // IEEE Transactions on Acoustics, Speech and Signal Processing. – 1987. – Vol. 5. – P. 404–406 (1987).
212. Rosendo Macias, J.A. Recursive formulation of short-time discrete trigonometric transforms / J.A Rosendo Macias, A. Gomez Exposito // IEEE Transactions on Circuits and Systems. II. – 1998. – Vol. 45. – P. 525-527.
213. Kober, V. Fast recursive algorithms for short-time discrete cosine transform / V. Kober, V., G. Cristobal // Electronics Letters. – 1999. – Vol. 35. – P. 1236-1238.

214. Xi, J. Computing running DCT's and DST's based on their second-order shift properties / J. Xi and J. F. Chiraro // IEEE Transactions on Circuits and Systems. I. – 2000. – Vol. 47. – P. 779-783.
215. Kober, V. Fast algorithms for the computation of sliding discrete sinusoidal transforms / V. Kober // IEEE Transactions on Signal Processing. – 2004. – Vol. 52. – P. 1704-1710.
216. Kelley, W. G. Difference Equations: An Introduction with Applications / W. G. Kelley and A. C. Peterson. – NY: Academic Press; 2 edition. 2000. – 403 P.
217. Mickens, R. Difference Equations / R. Mickens. – NY: Chapman and Hall/CRC. 1991. – P. 464.
218. Johnson, S.G. A modified split-radix FFT with fewer arithmetic operations/ S.G. Johnson, M. Frigo // IEEE Transactions on Signal Processing. – 2007. – Vol. 55. – P. 111-119.
219. Shao, X. Johnson Type-II/III DCT/DST algorithms with reduced number of arithmetic operations / X. Shao, S. G. Johnson // Signal Process. – 2008. – Vol. 88. – P. 1553-1564.
220. Wiener, N. Extrapolation, Interpolation and Smoothing of Stationary Time Series / N. Wiener. – N. Y.: MIT Press. 1964. – 171 P.
221. Мельниченко, А., Методы поиска изображений по визуальному подобию и детекции нечетких дубликатов изображений / Мельниченко, А. Гончаров // ЛММИИ на РОМИП-2009: Труды РОМИП (Петрозаводск, Россия Сентябрь 2009). – 2009. СПб: НУ ЦСИ. – С. 108-121.
222. Mojsilović, R. Matching and retrieval based on the vocabulary and 121 grammar of color patterns / R. Mojsilović, J. Kovačević, J. Hu, R. J. Safranek, S. K. Ganapathy // IEEE Transactions on Image Processing. – 2000. – Vol. 9. – P. 38-54.
223. Tamura, H. Textural features corresponding to visual perception / H. Tamura, S. Mori, T. Yamawaki // IEEE Transactions on. Systems, Man and Cybernetics. – 1978. – Vol. 8 (6). – P. 460-473.
224. Murala, S. Color and texture features for image indexing and retrieval / S. Murala, A.B. Gonde, R. P., Maheshwari // Proceedings of the IEEE International Advance

Computing Conference Advance Computing Conference, IACC (Patiala, India, March 2009). – 2009. – C. 1411-1416.

225. Stricker, M.A. Color indexing with weak spatial constraints / M. A. Stricker, A. Dimai // Proceedings of the SPIE Electronic Imaging: Science and Technology: Storage and Retrieval for Still Image and Video Databases IV (San Jose, USA, February 1996). – 1996. -Vol. 2670. – P. 29-40.

226. Squire, D. M. Content-based query of image databases, inspirations from text retrieval: inverted files, frequency-based weights and relevance feedback / D. M. Squire, W. Müller, H. Müller, J. Raki // In Pattern Recognition Letters. –1999. – P. 143-149.

227. Vidal-Calleja, T.A. Event-driven loop closure in multi-robot mapping / T. A. Vidal-Calleja, C. Berger, S. Lacroix // Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS (St Louis, United States October 2009). – 2009. – P. 1535-1540.

228. Angeli, A. Real-time visual loop-closure detection / A. Angeli, S. Doncieux, J. A. Meyer, D. Filliat // Proceedings of the IEEE International Conference on Robotics and Automation (Pasadena, CA, USA May 2008). – 2008. – P. 18421847.

229. Demirdjian, D. Avoiding the streetlight effect: Tracking by exploring likelihood modes / D. Demirdjian, L. Taycher, G. Shakhnarovich, K. Grauman, and T. Darrell // Proceedings of the IEEE International Conference on Computer Vision, , (Beijing, China, October 2005). – 2005. Vol. 1 – P. 357364.

230. Csurka, G. Visual categorization with bags of keypoints / G. Csurka, C. R. Dance, L. Fan, [et al] // Proceedings of the European Conference on Computer Vision, ECCV (Prague, Czech Republic, May 2004). – 2004. Vol. 1. – P. 1-22.

231. Botterill, T. "Bag-of-words-driven, single-camera simultaneous localization and mapping / T. Botterill, Tom, M. Steven, R. Green // Journal of Field Robotics. – 2011. – Vol. 28 (2). – P. 204-226.

232. Winn, J. Object categorization by learned universal visual dictionary / J. Winn, A. Criminisi, T. Minka // Proceedings IEEE International Conference on Computer Vision (Beijing, China, October 2005). – 2005. – Vol. 2. –P. 18001807.

233. Nister D. Scalable recognition with a vocabulary tree / D. Nister, H. Stewenius // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (New York, USA, June 2006). – 2006. – Vol. 2. – P. 2161-2168.
234. Zheng, L. Lp-norm idf for large scale image search / Zheng, L. [et al] / Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR (Portland, OR, USA, June 2013). – 2013. – P. 1626-1633.
235. Chen, D. M. Inverted Index Compression for Scalable Image Matching / D. M. Chen, S.T. Sam S., V. Chandrasekhar, G. Takacs, R. Vedantham, R. Grzeszczuk, B. Girod // Proceedings of the Data Compression Conference (Snowbird, UT, USA, March 2010). – 2010. – P. 1-10.
236. LeCun, Y. Convolutional networks for images, speech and time series / Y. LeCun // The Handbook of Brain Theory and Neural Networks. – Cambridge: MIT Press Cambridge, USA. 1996. – P. 255-258.
237. Mitra, V. Hybrid convolutional neural networks for articulatory and acoustic information based speech recognition / V. Mitra, G. Sivaraman, H. Nam, C. Espy-Wilson, E. Saltzman, M. Tiede // Journal Speech Communication archive. – 2017. – Vol. 89(C). – P. 103-112.
238. Wang, W. Generalized Autoencoder: A Neural Network Framework for Dimensionality Reduction / W. Wang, Y. Huang, Y. Wang, L. Wang // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, SVRP (Columbus, Ohio, USA, June 2014). – 2014. – .P. 490-497 (2014).
239. Rifai, S. Contractive auto-encoders: Explicit invariance during feature extraction / S. Rifai, P. Vincent, X. Muller, X. Glorot, Y. Bengio // Proceedings of the 28th International Conference on International Conference on Machine Learning (Bellevue, Washington, USA, June - July 2011). – 2011. – P. 833-840.
240. Forsyth, D. Computer Vision: A Modern Approach / D. Forsyth, J. Ponce. – Prentice Hall Professional Technical Reference, 2002.
241. LeCun, Y. Convolutional networks and applications in vision / Y. LeCun, K., Kavukvuoglu, C. Farabet // Proceedings of the IEEE International Symposium on Circuits and Systems (Paris, France, May –June 2010). – 2010. – P. 253–256.

242. Labeled Faces in the Wild, LFW (College of Information and Computer Sciences, University of Massachusetts Amherst). [Электронный ресурс]. Режим доступа: <http://vis-www.cs.umass.edu/lfw/>. (дата обращения: 01.12.2019).

243. Viola, P., Jones, M.J. Robust real-time face detection / P. Viola, M.J, Jones // International Journal of Computer Vision. – 2004. – Vol. 57 (2). – P. 137– 154.

244. Viola, P. Rapid Object Detection using a Boosted Cascade of Simple Features / P. Viola, M.J. // Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition , CVPR (Kauai, HI, USA, December 2001). – 2003. – P. 1-9.

245. Dalal, N. Histograms of oriented gradients for human detection / N. Dalal, B. Triggs // Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition , CVPR (San Diego, USA, June 2005). – 2005. – Vol. 1. – P. 886-893.

246. Squire, D. Content-based query of image databases, inspirations from text retrieval: inverted files, frequency-based weights and relevance feedback / D. Squire, W. Muller, H. Muller, J. Raki // Proceedings of the 11th Scandinavian Conference on Image Analysis (Kangerlussuaq, Greenland June 1999). – 1999. – P. 143.-149.

247. Kemelmacher-Shlizerman, I. The megaface benchmark: 1 million faces for recognition at scale / I. Kemelmacher-Shlizerman, S. Seitz, D. Miller, E. Brossard. – arXiv preprint arXiv:1512.00596. – 2015.

248. Labeled faces in the wild: A database for studying face recognition in unconstrained environments [Текст]: Technical Report: 07-49 / University of Massachusetts, Amherst; G. B. Huang, M. Ramesh, T. Berg, E. Learned-Miller. – Boston, 200. – Vol. 1 (2- 3) – 11 P.

**Приложение А. Свидетельства о государственной регистрации
программы для ЭВМ**

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО
о государственной регистрации программы для ЭВМ

№ 2015662478

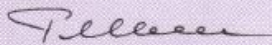
**Программная система восстановления изображений от
мультипликативных интерференций на основе
микросканирования камеры**

Правообладатель: *Федеральное государственное бюджетное
образовательное учреждение высшего профессионального
образования «Челябинский государственный университет» (RU)*

Авторы: *Вохминцев Александр Владиславович (RU),
Кобер Виталий Иванович (RU)*

Заявка № **2015619472**
Дата поступления **07 октября 2015 г.**
Дата государственной регистрации
в Реестре программ для ЭВМ **25 ноября 2015 г.**

Руководитель Федеральной службы
по интеллектуальной собственности

 **Г.П. Ивлиев**



РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2015662479

**Централизованная система дистанционной некооперативной
избирательной мультибиометрической аутентификации
личности**

Правообладатель: *Федеральное государственное бюджетное
образовательное учреждение высшего профессионального
образования «Челябинский государственный университет» (RU)*

Авторы: *Вохминцев Александр Владиславович (RU),
Ручай Алексей Николаевич (RU)*

Заявка № 2015619451

Дата поступления 07 октября 2015 г.

Дата государственной регистрации

в Реестре программ для ЭВМ 25 ноября 2015 г.



Руководитель Федеральной службы
по интеллектуальной собственности

Г.П. Ивлиев

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2015662752

Программная система обнаружения объектов на основе
дескриптора локальной гистограммы ориентированных
градиентов

Правообладатель: *Федеральное государственное бюджетное
образовательное учреждение высшего профессионального
образования «Челябинский государственный университет» (RU)*

Авторы: *Вохминцев Александр Владиславович (RU),
Мельников Андрей Витальевич (RU)*



Заявка № 2015619471

Дата поступления 08 октября 2015 г.

Дата государственной регистрации

в Реестре программ для ЭВМ 02 декабря 2015 г.

Руководитель Федеральной службы
по интеллектуальной собственности

Г.П. Ивлиев

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2019660919

**Адаптивная система одновременной навигации и
составления карты в неизвестном трехмерном пространстве
на основе динамического сопоставления изображений и
регистрации облаков точек**

Правообладатель: **Федеральное государственное бюджетное
образовательное учреждение высшего образования «Челябинский
государственный университет» (RU)**

Авторы: **Вохминцев Александр Владиславович (RU),
Кобер Виталий Иванович (RU)**

Заявка № 2019619711

Дата поступления 05 августа 2019 г.

Дата государственной регистрации

в Реестре программ для ЭВМ 15 августа 2019 г.

Руководитель Федеральной службы
по интеллектуальной собственности

Г.П. Ивлиев



РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2019660920

Комбинированный метод регистрации мультисенсорных данных на основе точного решения вариационной подзадачи итеративного алгоритма ближайших точек для аффинных и ортогональных преобразований

Правообладатель: **Федеральное государственное бюджетное образовательное учреждение высшего образования «Челябинский государственный университет» (RU)**

Авторы: **Вохминцев Александр Владиславович (RU),
Кобер Виталий Иванович (RU)**

Заявка № 2019619707

Дата поступления 05 августа 2019 г.

Дата государственной регистрации

в Реестре программ для ЭВМ 15 августа 2019 г.



Руководитель Федеральной службы
по интеллектуальной собственности

Г.П. Ивлиев

Приложение Б. Акты внедрения

НАУЧНО-ПРОИЗВОДСТВЕННОЕ ОБЪЕДИНЕНИЕ
АНДРОИДНАЯ ТЕХНИКА

АКТ

об использовании результатов диссертационной работы Вохминцева Александра Владиславовича «Методология решения проблемы одновременной навигации и построения карты на основе комбинирования визуальных и семантических характеристик окружающей среды», представленной на соискание ученой степени доктора технических наук

Настоящим актом подтверждаем, что результаты, полученные Вохминцевым Александром Владиславовичем в диссертационной работе на тему «Методология решения проблемы одновременной навигации и построения карты на основе комбинирования визуальных и семантических характеристик окружающей среды», были апробированы в научно-производственном объединении «Андроидная техника».

Предложенная автором работы методология для надежного точного решения задачи одновременной навигации и картографирования, основанная на комбинировании визуальных и семантических характеристик неизвестной окружающей среды, позволяет повысить степень автономности управления роботизированным комплексом и создавать системы SLAM для решения широкого класса прикладных 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. Предложенная автором трехмерная картографическая система основана на трех компонентах: пространственное совмещение последовательных RGB-кадров данных; определение положения в пространстве и обнаружение замыканий цикла; семантическая маркировка пространства.

Разработанная система SLAM прошла апробацию при решении задачи спасения людей в опасных и труднодоступных для человека инфраструктурных объектах. Данная работа была выполнена при частичной

НАУЧНО-ПРОИЗВОДСТВЕННОЕ ОБЪЕДИНЕНИЕ
АНДРОИДНАЯ ТЕХНИКА

поддержке Российского фонда фундаментальных исследований: проект № 18-37-20032 «Методы управления автономными техническими объектами на основе планирования траектории в среде с динамическими препятствиями» и проект № 16-08-00342 «Разработка системы одновременной навигации и составления карты в неизвестном пространстве на основе динамического сопоставления изображений и регистрации облаков точек».

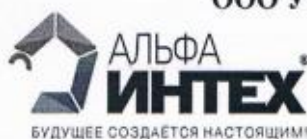
Исполнительный директор



Е.А. Дудоров

1.06.2020

АО «НПО «Андроидная техника»
109518, г. Москва, ул. Грайворонская, д. 23
ОГРН 1097746741894, ИНН/КПП 7723738378/772201001
тел.: +7 (495) 226 02 99, +7 (495) 226 12 99, +7 (495) 226 14 99
факс: +7 (3519) 39 52 58
e-mail: info@rusandroid.com, www.npo-at.com



ООО Уральский Робототехнологический Центр «Альфа-Интех»

454038, г. Челябинск, ул. Ш. Металлургов, 27П/1, оф.10

ИНН 7460017486 КПП 746001001 БИК 047511711

р/с 40702810207130005689 в ОАО «Челиндбанк»

Металлургический филиал к/с 30101810400000000711

СПРАВКА ОБ ИСПОЛЬЗОВАНИИ

результатов диссертационной работы Вохминцева Александра Владиславовича на тему «Методология решения проблемы одновременной навигации и построения карты на основе комбинирования визуальных и семантических характеристик окружающей среды», представленной на соискание ученой степени доктора технических наук

Настоящим документом подтверждается, что основные положения и результаты диссертационного исследования Вохминцева А.В. внедрены и апробированы в научной и проектной деятельности ООО Уральского робототехнического центра «Альфа-Интех».

Предложенная в диссертационной работе комбинированная методология для решения задачи одновременной навигации и картографирования позволяет создавать системы SLAM (Simultaneous localization and mapping), рекомендованные к использованию в промышленных роботах и робототехнических комплексах, а также в различных приложениях робототехники с дополненной реальностью.

Методы и алгоритмы методологии:

- алгоритм поиска особых точек на изображениях,
- методы решения вариационной задачи для регистрации трехмерных данных,
- метод решения задачи навигации и определения положения мобильной платформы, основанный на комбинировании визуальных и семантических характеристик неизвестной окружающей среды и расширенной Калмановской фильтрации (EKF)

были использованы в аппаратно-программных приложениях компании.

Применение мультисенсорной и семантической информации на всех этапах решения проблемы SLAM для повышения точности и сходимости указанных методов и алгоритмов методологии позволяет осуществлять одновременную навигацию и картографирование для крупномасштабных сцен в реальном масштабе времени, в том числе для сцен с неравномерным освещением.

Разработанная система SLAM прошла апробацию в следующих проектах компании:

- программно-аппаратная платформа для дистанционного управления и программирования роботизированных комплексов различного назначения;
- системе машинного зрения робототехнических комплексов.

Результаты диссертационной работы планируются к использованию компанией при выполнении НИОКР по теме «Программно-аппаратная платформа для обеспечения совместной работы распределенных групп пользователей в режиме реального времени при ассистировании операторам робототехнических систем и дистанционного программирования роботов-манипуляторов с визуализацией данных методами дополненной реальности».

24 июля 2020 г.

Генеральный директор



А.А. Перепелица

УТВЕРЖДАЮ:

Директор Югорского научно-исследовательского института информационных технологий
д.т.н., профессор



А.В. Мельников

«24» июля 2020 г.

Акт о применении результатов

диссертационной работы Вохминцева Александра Владиславовича на тему «Методология решения проблемы одновременной навигации и построения карты на основе комбинирования визуальных и семантических характеристик окружающей среды», представленной на соискание ученой степени доктора технических наук

Результаты диссертационной работы, полученные Вохминцевым Александром Владиславовичем, были апробированы в Югорском научно-исследовательском институте информационных технологий при решении проблемы картографирования фильтрующих каналов в больших объемах сейсмических данных.

Комбинированные методы регистрации данных использованы при решении задачи локализации зон трещиноватости и создании геопространственной технологии и геоинформационной платформы для освоения глубоководных минеральных ресурсов, в частности:

- для автоматизации выделения зон трещиноватости на 2D-изображениях (сейсмический разрез) и
- для автоматизации выделения зон трещиноватости на 3D моделях (сейсмический куб).

Комбинированный метод решения вариационной задачи точка-плоскость для класса ортогональных преобразований использован в задаче интерпретации при разметке данных, которая сводится к задаче сравнения трехмерных поверхностей.

Комбинированные методы регистрации данных позволяют:

- автоматизировать процесс интерпретации данных сейсморазведки, связанный с выделением геологических слоев, разломов и горизонтов на сейсмических разрезах и кубах;
- решить проблему неоднозначности интерпретации данных при обработке данных сейсморазведки.

Для верификации метода регистрации данных был использован эталонный набор данных OpenSeismicRepository, который содержит информацию о геологических породах в районе акватории Северного моря и 3D данные сейсморазведочной информации в объеме 70 разрезов волнового поля с территории Ямало-Ненецкого автономного округа.

Руководитель центра
космических услуг, к.т.н.



Кочергин Г.А.

УТВЕРЖДАЮ

Генеральный директор научно-
производственного центра
«Интеллектуальные транспортные
системы»

д.т.н., профессор Михеева Т.И.

«05» 08 2020 г.

**Акт об использовании результатов диссертационной работы Вохминцева
Александра Владиславовича на тему «Методология решения проблемы
одновременной навигации и построения карты на основе комбинирования
визуальных и семантических характеристик окружающей среды»**

Комиссия в составе зам. генерального директора Алексеева С.А. и начальника информационного отдела Хаева А.С. составили настоящий акт о том, что *основные результаты диссертационной работы* заведующего научно-исследовательской лабораторией интеллектуальные информационные технологии и системы ФГБОУ ВО «Челябинский государственный университет» Вохминцева А.В. на соискание ученой степени доктора технических наук по специальности 05.13.01 – «Системный анализ, управление и обработка информации (в информационных и технических системах)» *были апробированы* в научно-производственном центре «Интеллектуальные транспортные системы», а именно:

- алгоритм для поиска и сопоставления особых точек на двумерных изображениях и
- точные комбинированные методы реконструкции контекстуально сложных крупномасштабных трехмерных сцен на основе итеративного алгоритма ближайших точек.

Результаты диссертации были использованы в системе визуализации информации геоинформационной системы ITSGIS, которая предназначена для автоматизации работ, выполняющих функции учета объектов городской и транспортной инфраструктуры.

Предложенные в работе методы и алгоритмы позволили существенно повысить эффективность и точность процесса обработки информации в подсистеме визуализации данных при построении двухмерных и трехмерных моделей окружающего пространства.

Зам. генерального директора

Алексеев С.А.

Начальник информационного отдела

Хаев А.С.





МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное
образовательное учреждение
высшего образования

«Челябинский государственный университет»
(ФГБОУ ВО «ЧелГУ»)

ул. Братьев Кашириных, 129, г. Челябинск, 454001
тел. (351) 799-71-01, факс: (351) 742-09-25
E-mail: odou@csu.ru; http://www.csu.ru
ОКПО 05121292, ОГРН 1027402324905,
ИНН/КПП 7447012841/744701001

УТВЕРЖДАЮ

Ректор ФГБОУ ВО «Челябинский
государственный университет»

С.В. Таскаев
«15» ср 2022 г.

№ _____

На № _____ от _____

АКТ

Использования в учебном процессе результатов диссертационной работы
«Методология решения проблемы одновременной навигации и построения карты
на основе комбинирования визуальных и семантических характеристик
окружающей среды» заведующего научно-исследовательской лабораторией
интеллектуальные информационные технологии и системы
ФГБОУ ВО «Челябинский государственный университет»
Вохминцева Александра Владиславовича,
представленной на соискание ученой степени доктора технических наук

Подтверждаем, что результаты, полученные Вохминцевым А.В. в
диссертационной работе на тему «Методология решения проблемы
одновременной навигации и построения карты на основе комбинирования
визуальных и семантических характеристик окружающей среды» были
апробированы в ФГБОУ ВО «Челябинский государственный университет»
(далее – ЧелГУ).

В учебном процессе ЧелГУ при обучении студентов института
информационных технологий по направлениям 09.03.04 Программная
инженерия и 02.04.02 Фундаментальная информатика и информационные
технологии используются следующие результаты диссертационной работы:

- алгоритм представления окрестности особых точек в виде вектора признаков на основе рекурсивного вычисления гистограмм направленных градиентов,
- комбинированные методы решения вариационной задачи с метриками точка-точка и точка-плоскость на основе метода ИСР для аффинных и ортогональных преобразований,
- метод решения задачи навигации и определения положения мобильной платформы на сцене, основанный на комбинировании визуальных и се-

004330

математических характеристик неизвестной окружающей среды и расширенной Калмановской фильтрации,

- подход к оптимизации графа положений мобильной платформы и обнаружению замыканий цикла на основе метода «корзина слов» и двумерного дескриптора на основе гистограммы ориентированных градиентов,
- прототип системы одновременной навигации и построения карты.

Перечисленные результаты используются в лекционных курсах и на практических занятиях по следующим дисциплинам:

- «Основы электроники и робототехники»;
- «Системный анализ и моделирование»;
- «Машинное обучение и анализ данных».

Указанные результаты также планируются к использованию в новом курсе «Робототехнические системы и комплексы» для студентов направления 02.04.02 Фундаментальная информатика и информационные технологии (магистратура).

Директор института
информационных технологий



Ю.В. Петриченко