**Министерство науки и высшего образования Российской Федерации**

**федеральное государственное автономное образовательное учреждение высшего образования**

**«Национальный исследовательский университет ИТМО»**

**(Университет ИТМО)**

Факультет: ПИиКТ

Образовательная программа: Технология интернета вещей

Направление подготовки (специальность): Программная инженерия

О Т Ч Е Т

по

Научно-исследовательской работе

Тема задания: *Применение методов трекинга и навигации с участием камеры Intel RealSense Tracking Camera T265*

Обучающийся: *Криворотко Алексей Александрович*

Согласовано: *Руководитель практики: Шматков Владислав Николаевич, кандидат технических наук, доцент факультета программной инженерии и компьютерной техники*

Практика пройдена с оценкой \_\_\_\_

Дата \_\_\_\_\_\_\_

Санкт-Петербург

2024

**Содержание**

[Обозначения и сокращения 4](#_Toc158100504)

[Введение 5](#_Toc158100505)

[Обзор литературы 6](#_Toc158100506)

[1.1 Evaluation of the Intel RealSense T265 for tracking natural human head motion 6](#_Toc158100507)

[1.2 Computer Vision for Sports: Current Applications and Research Topics 7](#_Toc158100508)

[1.3 Real-Time Tracking of Single and Multiple Objects from Depth-Colour Imagery Using 3D Signed Distance Functions 8](#_Toc158100509)

[1.4 IMU-based human activity recognition and payload classification for low-back exoskeletons 9](#_Toc158100510)

[1.5 Review of visual odometry: types, approaches, challenges, and applications 10](#_Toc158100511)

[1.6 The Surprising Effectiveness of Visual Odometry Techniques for Embodied PointGoal Navigation 10](#_Toc158100512)

[1.7 On Autonomous Spatial Exploration with Small Hexapod Walking Robot using Tracking Camera Intel RealSense T265 11](#_Toc158100513)

[Обзор камеры Intel RealSense Tracking Camera T265 13](#_Toc158100514)

[2.1 Описание камеры 13](#_Toc158100515)

[2.2 Технические характеристики 14](#_Toc158100516)

[2.3 Принцип работы 15](#_Toc158100517)

[2.4 Применение 16](#_Toc158100518)

[Методы трекинга и навигации с использованием камеры Intel RealSense T265 18](#_Toc158100519)

[3.1 IMU Fusion 18](#_Toc158100520)

[3.1.1 Акселерометр и гироскоп 18](#_Toc158100521)

[3.1.2 Вклад IMU в SLAM 21](#_Toc158100522)

[3.2 Visual Odometry 21](#_Toc158100523)

[3.2.1 Алгоритмы отслеживания особенностей 22](#_Toc158100524)

[3.2.2 Вклад Visual Odometry в SLAM 23](#_Toc158100525)

[3.3 SLAM 24](#_Toc158100526)

[Реализация трекинга и навигации 28](#_Toc158100527)

[4.1 Установка окружения 28](#_Toc158100528)

[4.2 Изучение SDK 2.0 29](#_Toc158100529)

[4.3 Написание скриптов на pyrealsense2 30](#_Toc158100530)

[Заключение 34](#_Toc158100531)

[Список использованных источников 35](#_Toc158100532)

# Обозначения и сокращения

В научно-исследовательской работе встречаются следующие термины с соответствующими определениями:

ПО – программное обеспечение.

IMU (Inertial Measurement Unit) – это электронное устройство, которое измеряет и сообщает о скорости, ориентации и гравитационных силах, используя комбинацию акселерометров и гироскопов.

VO (Visual Odometry) – визуальная одометрия, метод оценки перемещения по изображениям, полученным с камеры.

FAST (Features from Accelerated Segment Test) – алгоритм для обнаружения углов на изображении, используется для определения ключевых точек, быстрый и эффективный в работе.

SIFT (Scale-Invariant Feature Transform) – алгоритм в области компьютерного зрения для обнаружения и описания локальных особенностей на изображениях.

SURF (Speeded Up Robust Features) –алгоритм, используемый для обнаружения и описания локальных особенностей. Является усовершенствованием SIFT, оптимизированным для повышения скорости обработки при сохранении схожей производительности.

ORB (Oriented FAST and Rotated BRIEF) – алгоритм, сочетающий в себе быстрое обнаружение особенностей с помощью FAST и эффективное их описание с использованием BRIEF. Предлагает быстрое и эффективное решение для задач трекинга и распознавания.

SLAM (Simultaneous Localization and Mapping) – одновременная локализация и картографирование, комплексный процесс, при котором устройство одновременно строит карту окружающей среды и определяет свое местоположение в этой среде.

# Введение

В современном мире технологии компьютерного зрения становятся основополагающими для развития множества технических дисциплин, включая робототехнику и автономные транспортные системы. Важным аспектом в этом направлении является разработка и использование камер трекинга, которые могут в реальном времени обрабатывать пространственные данные. Intel RealSense Tracking Camera T265 выделяется среди подобных технологий благодаря своим способностям по трехмерному восприятию пространства, что делает её важным инструментом для решения задач трекинга и навигации.

Целью данной научно-исследовательской работы является изучение возможностей камеры Intel RealSense T265 в контексте трекинга и навигации в трехмерном пространстве. Мы проанализируем актуальные научные работы по данной теме, рассмотрим теоретические основы технологий SLAM, Visual Odometry и IMU Fusion, и оценим их применение вместе с камерой T265. Основная задача исследования — реализовать методы трекинга и навигации, используя эту камеру, и продемонстрировать их практическую эффективность.

Мы сосредоточимся на техническом анализе и практической реализации, стремясь обеспечить конкретные результаты, которые могут быть применены в реальных системах трекинга и навигации. Это исследование направлено на достижение понимания, как камера T265 может улучшить точность и надежность трекинга в различных приложениях, от автоматизированных роботов до систем автономного вождения и интерактивных приложений дополненной реальности.

# **Обзор литературы**

В настоящее время, технологии трекинга и навигации играют ключевую роль в развитии современных систем взаимодействия между человеком и машиной, автономных транспортных средств, робототехники и многих других областей. Одним из передовых решений в этой сфере является камера Intel RealSense Tracking Camera T265, которая обеспечивает высокую точность трекинга и навигации в реальном времени благодаря использованию передовых алгоритмов компьютерного зрения и интегрированных датчиков глубины. Данный обзор посвящен анализу существующих методов трекинга и навигации, применяемых с использованием данной камеры, а также исследованию их практического применения в различных сферах.

Цель данного обзора состоит в том, чтобы представить разнообразие подходов и методов в области трекинга и навигации, с акцентом на использование камеры Intel RealSense T265, а также рассмотреть исследования, расширяющие рамки этих технологий через применение компьютерного зрения и машинного обучения в других сферах.

Обзор включает в себя анализ семи научных работ, каждая из которых вносит вклад в понимание возможностей и ограничений камеры Intel RealSense T265, а также технологий, тесно связанных с ней.

## 1.1 Evaluation of the Intel RealSense T265 for tracking natural human head motion

В статье «Evaluation of the Intel RealSense T265 for tracking natural human head motion» [1] исследуется применение камеры Intel RealSense T265 для отслеживания движений головы человека. Авторы ставят задачу оценить точность и надёжность камеры в реальных условиях. Основное внимание уделяется сравнению данных, полученных с помощью камеры T265, с результатами, полученными с использованием других методов трекинга. Особенностью этого исследования является применение методов оптического трекинга для верификации данных, собранных с помощью T265.

Исследователи применяют комплексный подход, включающий использование как камеры RealSense, так и внешних систем трекинга для получения более объективной картины эффективности камеры. В статье подробно описывается методология экспериментов, включая настройку оборудования и процедуры обработки данных.

Результаты исследования описывают потенциал использования камеры Intel RealSense T265 в приложениях, связанных с трекингом движений головы человека, но также указывают на некоторые ограничения и предлагают направления для дальнейшего улучшения. Статья важна для специалистов, работающих в области разработки интерфейсов человеко-машинного взаимодействия и виртуальной реальности.

## 1.2 Computer Vision for Sports: Current Applications and Research Topics

Статья "Computer Vision for Sports: Current Applications and Research Topics" [2] описывает, как компьютерное зрение совершенствует спортивную индустрию, от улучшения производительности спортсменов до повышения уровня вовлеченности болельщиков. Она демонстрирует, как автоматизированный анализ видеоданных способствует более точному и объективному изучению игровых действий, движений игроков и их взаимодействий в игре. С помощью примеров из различных видов спорта, авторы обсуждают, как технологии компьютерного зрения применяются для автоматического отслеживания движений, анализа тактики и стратегии команд, а также для улучшения безопасности участников.

Особое внимание в статье уделяется инновационным методам обучения и анализа данных, которые позволяют улучшить тренировочные процессы и стратегии игры. Также рассматривается, как компьютерное зрение способствует разработке новых инструментов для судей и тренеров, позволяя им получать более точную обратную связь в реальном времени. Статья подчеркивает, что совместная работа исследователей, разработчиков и спортивных организаций является ключом к дальнейшему прогрессу в этой области.

В заключение, статья подводит итоги текущего состояния исследований в области компьютерного зрения в спорте и очерчивает будущие направления для развития. Это исследование является ценным ресурсом для разработчиков, в области компьютерного зрения, стремящимся интегрировать новейшие технологии в свою работу, так как разбирает эти технологии на примере спорта.

## 1.3 Real-Time Tracking of Single and Multiple Objects from Depth-Colour Imagery Using 3D Signed Distance Functions

Статья «Real-Time Tracking of Single and Multiple Objects from Depth-Colour Imagery Using 3D Signed Distance Functions» [3] детально рассматривает разработку и применение методики для отслеживания одиночных и множественных объектов в реальном времени с использованием данных глубины и цвета. Авторы представляют подход, основанный на 3D функциях знакового расстояния (SDF), для создания вероятностных моделей, которые могут эффективно интерпретировать и анализировать сложные видеоданные. Этот подход позволяет системе учитывать информацию о форме и внешности объектов, значительно повышая точность и надёжность процесса отслеживания.

Исследователи подробно описывают методологию, демонстрируя, как их система может адаптироваться к динамичным изменениям в среде и взаимодействиям между объектами, обеспечивая при этом высокую скорость обработки. Особое внимание уделяется проблемам, связанным с окклюзией и перекрытием объектов, и предлагаются эффективные стратегии для их преодоления. Авторы также проводят сравнение своего подхода с другими существующими методами, подчеркивая преимущества использования SDF для комплексного анализа видеоданных в различных условиях. Результаты экспериментов, представленные в статье, демонстрируют значительное улучшение в точности и скорости отслеживания.

Подводя итоги, данная статья предлагает взгляд на будущее технологий компьютерного зрения, где точное и эффективное отслеживание объектов в реальном времени становится возможным благодаря инновационным методам и подходам. Это исследование может помочь разработчикам систем компьютерного зрения, сталкивающимся с проблемами, связанными с окклюзией и перекрытием объектов, так как она предлагает варианты решения этих задач.

## 1.4 IMU-based human activity recognition and payload classification for low-back exoskeletons

Статья «IMU-based human activity recognition and payload classification for low-back exoskeletons» [4] представляет результаты исследования, направленного на разработку и оценку системы распознавания деятельности человека и классификации нагрузки, основанной на использовании инерциальных измерительных устройств (IMU). Эта работа ориентирована на улучшение функциональности и эффективности низкопрофильных экзоскелетов для поясницы, применяемых в промышленных условиях для снижения риска развития мускулоскелетных заболеваний среди рабочих. Авторы подробно описывают алгоритмы машинного обучения и глубокого обучения, разработанные для точного распознавания физических действий пользователя и классификации массы поднимаемых грузов.

Исследование демонстрирует, как данные с IMU можно использовать для генерации детальных и точных моделей активности, которые способны уловить даже незначительные изменения в движении и позиции тела, делая возможным реализацию реального времени адаптации функций экзоскелета к текущим операциям пользователя. Авторы также обсуждают потенциальные применения разработанной системы за пределами индустриального использования, включая спортивную тренировку и реабилитацию после травм.

Таким образом, статья подчеркивает значительный потенциал применения IMU в разработке умных экзоскелетов, способствующих снижению риска профессиональных заболеваний и улучшению качества жизни рабочих. Данная статья может быть полезна разработчикам для изучения работы инерциальных измерительных устройств (IMU) и применения их в алгоритмах машинного и глубокого обучения.

## 1.5 Review of visual odometry: types, approaches, challenges, and applications

Статья "Review of visual odometry: types, approaches, challenges, and applications" [5] представляет собой глубокий аналитический обзор, который исследует ключевые аспекты и достижения в области визуальной одометрии (VO). В ней обсуждаются различные типы VO, включая моновизуальную и стереовизуальную одометрию, а также основанные на глубинном зрении методы. Авторы статьи подробно рассматривают подходы к реализации VO, включая комбинированные подходы, которые стремятся улучшить точность и устойчивость системы к изменениям окружающей среды и различным условиям освещения. Особое внимание уделяется вызовам, с которыми сталкиваются разработчики при внедрении VO в практические приложения, включая проблемы, связанные с динамическими изменениями в сцене, ограничениями оборудования и необходимостью обеспечения высокой вычислительной эффективности. Статья также освещает широкий спектр приложений VO, от автономных автомобилей и дронов до виртуальной и дополненной реальности, указывая на обширность этой области.

Подводя итоги, этот обзор станет ценным ресурсом для исследователей, инженеров и разработчиков, работающих в области автономных систем, робототехники и компьютерного зрения, предоставляя им подробное понимание текущего состояния технологии VO, ее возможностей и ограничений, а также перспектив для будущих разработок.

## 1.6 The Surprising Effectiveness of Visual Odometry Techniques for Embodied PointGoal Navigation

Статья «The Surprising Effectiveness of Visual Odometry Techniques for Embodied PointGoal Navigation» [6] детализирует исследование, демонстрирующее значительные преимущества использования техник визуальной одометрии (VO) в задачах навигации с ограничениями, типичными для искусственного интеллекта. Исследователи фокусируются на применении VO для улучшения точности и эффективности навигационных алгоритмов в условиях, где доступ к традиционным средствам локализации, таким как GPS или компас, ограничен или недоступен. В статье подробно рассматриваются методологические подходы к интеграции VO в системы навигации, а также проводится сравнение с другими современными методами в контексте точности и скорости выполнения задачи навигации.

Через серию экспериментов, проведенных в симуляционной среде, авторы демонстрируют, что VO может существенно улучшить способность алгоритмов точно определять местоположение и ориентацию робота в пространстве, что, в свою очередь, ведет к более эффективному и надежному достижению целевых точек. Эти результаты подчеркивают потенциал VO как мощного инструмента для разработки более продвинутых систем автономной навигации.

В заключение, статья предлагает новые перспективы для развития области воплощенного AI и автономной навигации, выделяя VO как ключевой элемент для создания эффективных и адаптивных навигационных систем, способных функционировать в сложных и динамически изменяющихся средах. Данная статья может быть полезна для разработчиков в области робототехники и автономных систем, работающих над улучшением алгоритмов навигации и локализации без использования GPS. Также она представляет интерес для специалистов в области искусственного интеллекта, занимающихся разработкой систем, которые должны эффективно ориентироваться в сложных средах.

## 1.7 On Autonomous Spatial Exploration with Small Hexapod Walking Robot using Tracking Camera Intel RealSense T265

Статья «On Autonomous Spatial Exploration with Small Hexapod Walking Robot using Tracking Camera Intel RealSense T265» [7] посвящена разработке и испытанию маленького робота, оснащенного камерой Intel RealSense T265, для автономного исследования пространства. Основная цель работы заключалась в демонстрации возможностей робота по самостоятельному картографированию и навигации в неизвестной среде без вмешательства человека. Авторы подробно описывают технические аспекты робота, включая его конструкцию, системы визуализации и алгоритмы обработки данных, которые позволяют ему эффективно перемещаться и создавать трехмерные карты окружающей среды.

В статье также рассматриваются вызовы, связанные с автономным исследованием малыми роботизированными платформами, такими как ограничения по размеру, весу и вычислительной мощности. Исследователи демонстрируют, как эти проблемы были преодолены с помощью инновационного подхода оптимизации алгоритмов.

Результаты экспериментов, представленные в статье, подтверждают, что разработанный робот способен эффективно справляться с задачей автономного пространственного исследования, открывая новые перспективы для применения подобных технологий в различных областях, например для поисково-спасательных операций.

Таким образом, данная статья представляет значительный интерес для разработчиков в области робототехники и автономных систем, предлагая новые идеи и подходы для создания малогабаритных роботизированных платформ, способных к эффективному исследованию и картографированию пространства.

# Обзор камеры Intel RealSense Tracking Camera T265

Перед тем как приступить к практической работе с камерой Intel RealSense Tracking Camera T265, для начала необходимо провести изучение её ключевых особенностей, технических характеристик и основных принципов работы. Данный этап поможет нам оценить потенциал данного устройства и выявить его возможности для решения конкретных задач в различных областях применения.

## 2.1 Описание камеры

Intel RealSense Tracking Camera T265 является высокотехнологичным устройством для визуального трекинга, специально разработанным для обеспечения высокоточного трехмерного позиционирования и ориентации в реальном времени. Это достигается за счет использования двух рыбьеглазных камер, формирующих стереопару, что позволяет устройству создавать детализированные глубинные карты окружающей среды. Эта особенность делает камеру отличным инструментом в ряде приложений, начиная от навигации автономных роботов и заканчивая разработкой интерактивных систем дополненной (AR) и виртуальной реальности (VR), где критически важны точность определения положения и ориентации устройства в пространстве без потребности в дополнительных внешних трекинговых системах. Камера изображена на рисунке 1.



Рисунок 1 – Камера Intel RealSense Tracking Camera T265

Ключевые преимущества Intel RealSense Tracking Camera T265 обусловлены несколькими факторами. Во-первых, использование технологии визуальной одометрии на основе одновременного картирования и локализации (V-SLAM) позволяет достигать высокой точности трекинга с минимальной задержкой, что критически важно для приложений, требующих реакции в реальном времени, например, для навигации дронов или автономных транспортных средств.

Вторым значительным преимуществом является независимость камеры от внешних систем трекинга. Это делает T265 идеальным выбором для мобильных и автономных приложений, где использование внешних систем трекинга может быть невозможным или неэффективным.

Третье преимущество – легкость интеграции. Благодаря своим компактным размерам и простому интерфейсу подключения через USB 3.1 Gen 1 Type-C, камера T265 может быть легко встроена в различные устройства и системы, не требуя сложной аппаратной интеграции или дополнительных адаптеров.

Четвёртое преимущество – многообразие применения камеры Intel RealSense Tracking Camera T265 является одним из ее главных достоинств. Высокая производительность и точность трекинга делают ее подходящей для широкого спектра приложений — от автономных роботизированных систем и беспилотных летательных аппаратов до интерактивных инсталляций AR/VR и различных исследовательских проектов. Эта универсальность позволяет разработчикам и инженерам выбирать T265 как надежное решение для решения задач трекинга и навигации в самых разнообразных областях.

## 2.2 Технические характеристики

К основным характеристикам камеры можно отнести [8]:

* Процессор: Intel Movidius Myriad 2.
* Две рыбьеглазных камеры: 848x800 пикселей на каждый глаз.
* Частота кадров: до 30 кадров в секунду.
* Поле зрения: около 163° по горизонтали и 85° по вертикали.
* Датчики: встроенные акселерометр и гироскоп.
* Интерфейс: USB 3.1 Gen 1 Type-C для подключения к устройствам.
* Совместимость: поддержка работы с широким спектром операционных систем, включая Windows, Linux.
* Вес: легкая и компактная конструкция для легкой интеграции.

## 2.3 Принцип работы

Принцип работы Intel RealSense Tracking Camera T265 основывается на технологии визуального одометрического трекинга (V-SLAM), которая обеспечивает высокую точность определения положения и ориентации камеры в пространстве.

V-SLAM комбинирует данные с камер и инерционных измерительных устройств (IMU), таких как акселерометр и гироскоп, для точного трекинга движения в трехмерном пространстве без необходимости в GPS или других внешних системах ориентации. Данную технологию можно условно разделить на три этапа.

Первый этап – это сбор данных:

* Стереопара: T265 оснащена двумя камерами, расположенными на определенном расстоянии друг от друга. Это позволяет камере захватывать стереоскопические изображения, необходимые для создания трехмерных карт окружающей среды.
* Инерционные датчики: акселерометр и гироскоп измеряют линейное ускорение и угловые скорости соответственно, предоставляя информацию о движении и ориентации устройства в пространстве.

Второй этап — это обработка данных:

* Анализ изображений: алгоритмы V-SLAM анализируют последовательные изображения с камер для определения изменений в положении и ориентации устройства, сравнивая особенности (такие как края, углы и другие характерные точки) на разных кадрах.
* Интеграция данных с IMU: данные с акселерометра и гироскопа интегрируются с визуальной информацией для коррекции ошибок и увеличения точности трекинга. Это позволяет системе поддерживать высокую точность даже в условиях низкой видимости или при быстром перемещении.

Заключительный этап — это создание карты и трекинг:

* Создание карты окружающей среды: система использует собранные данные для создания детализированной трехмерной карты пространства, где находится устройство. Эта карта включает в себя информацию о местоположении и ориентации объектов вокруг.
* Отслеживание перемещений: на основе созданной карты и постоянно обновляемых данных с камер и IMU, система может точно отслеживать движение и изменения ориентации устройства в пространстве, позволяя ему ориентироваться и перемещаться с высокой степенью автономности.

## 2.4 Применение

Камера T265 идеально подходит для широкого спектра приложений, где требуется точное трехмерное позиционирование и трекинг. Среди них:

* Автономные роботы и дроны: Intel RealSense T265 подходит для автономных систем, таких как роботы и дроны, предоставляя им способность самостоятельно перемещаться в сложной и динамичной среде. Используя V-SLAM, камера обеспечивает точное определение положения и ориентации без необходимости GPS или внешних маяков. Это позволяет роботам и дронам избегать препятствий, выполнять задачи поиска и спасения, а также обеспечивать доставку в условиях, где использование традиционных навигационных систем ограничено или невозможно.
* Дополненная и виртуальная реальность: в области AR и VR камера T265 улучшить опыт пользователя. Трекинг движений пользователя с высокой точностью позволяет улучшить взаимодействие с виртуальными объектами и средами, делая их более интуитивно понятными и естественными.
* Промышленные и исследовательские приложения: в промышленности и научных исследованиях камера T265 может служить для создания точных трехмерных карт рабочих пространств или мест проведения экспериментов. Это особенно полезно для задач, требующих автономного патрулирования, инспекции или сбора данных в условиях, где использование человека может быть опасным или непрактичным.
* Навигационные системы для людей с ограниченными возможностями: использование T265 может также быть расширено на разработку навигационных и помощников для людей с ограниченными возможностями, обеспечивая им большую самостоятельность и безопасность при перемещении в неизвестных или сложных для восприятия средах.
* И многие другие области применения, связанные с компьютерным зрением.

# Методы трекинга и навигации с использованием камеры Intel RealSense T265

Для того чтобы реализовать трекинг и навигацию с помощью камеры Intel Realsense T265 нам необходимо разобраться в технологиях участвующих в этом процессе.

В данном исследовании осуществляется анализ инерциальных измерительных устройств (IMU), изучение технологий визуальной одометрии (VO) и алгоритмов Simultaneous Localization And Mapping (SLAM) как ключевых элементов в системах навигации и отслеживания. Целью является понимание их взаимодействия и вклада в повышение точности и надежности навигационных систем.

## 3.1 IMU Fusion

Инерциальное измерительное устройство (IMU) в контексте Intel RealSense T265, играет фундаментальную роль в обеспечении точности и надежности навигации. Это устройство, состоящее из акселерометра и гироскопа, предоставляет данные о движении и ориентации устройства.

### 3.1.1 Акселерометр и гироскоп

Акселерометр и гироскоп в IMU работают на основе фундаментальных физических принципов, и их функционирование можно описать с помощью основных физических формул.

**Акселерометр**

Акселерометр измеряет линейное ускорение в трех измерениях, обеспечивая данные о скорости и направлении движения [9]. На рисунке 2 изображена схема простейшего акселерометра.

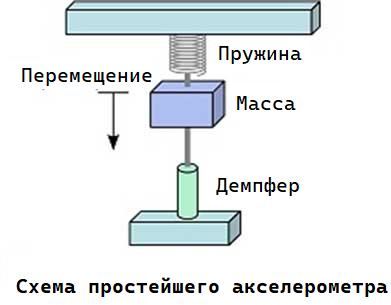


Рисунок 2 – Схема простейшего акселерометра

Ускорение можно выразить через вторую производную по времени от перемещения или через первую производную по времени от скорости:

|  |  |
| --- | --- |
| где | – ускорение;  – перемещение;  – скорость;  – время. |

В трехмерном пространстве это ускорение измеряется вдоль трех осей (X, Y, Z), что позволяет определить полное векторное ускорение объекта.

Формула для полного векторного ускорения выглядит следующим образом:

|  |  |
| --- | --- |
| где | – вектор полного ускорения;  – компоненты ускорения вдоль осей , и соответственно;  – единичные векторы в направлениях осей , и . |

Каждая из компонент определяется как производная скорости по времени в соответствующем направлении:

|  |  |
| --- | --- |
| где | – компоненты скорости вдоль осей , и . |

**Гироскоп**

Гироскоп отслеживает угловую скорость, что критически важно для определения ориентации устройства [10]. Схема простейшего гироскопа изображена на рисунке 3.

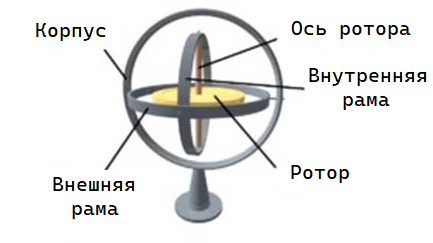


Рисунок 3 – Схема простейшего гироскопа

Гироскоп измеряет угловую скорость , которую можно выразить через изменение угла поворота за время:

|  |  |
| --- | --- |
| где | – угловая скорость;  – угол поворота;  – время. |

Угловая скорость измеряется в радианах в секунду (рад/с) и указывает, насколько быстро объект вращается вокруг оси.

**Слияние данных (Sensor Fusion)**

Один из популярных методов слияния данных – это использование фильтра Калмана. Он позволяет объединять данные с акселерометра и гироскопа для получения более точной оценки положения и ориентации. Этот фильтр работает на основе принципа рекурсивной оценки состояния системы, что позволяет ему обрабатывать входящие данные в реальном времени. Фильтр Калмана оптимально объединяет новую информацию с уже имеющейся оценкой состояния, минимизируя влияние шума и ошибок измерения. Это достигается путем вычисления прогнозируемого состояния на основе предыдущих измерений и корректировки этого состояния с учетом новых входных данных.

Эти формулы представляют основу для работы акселерометра и гироскопа в IMU. Данные с этих датчиков затем интегрируются и обрабатываются для создания полной картины движения и ориентации устройства в пространстве.

### 3.1.2 Вклад IMU в SLAM

Инерциально измерительные устройства (IMU) играют большую роль в системах SLAM. Они помогают системам SLAM, когда обычные камеры и датчики не справились бы со своей задачей, например, при плохом свете или когда объекты движутся быстро. Благодаря IMU, система SLAM может быстро реагировать на изменения в движении и точно следить за тем, где находится объект, даже если условия меняются.

В камерах Intel RealSense T265 IMU дают точные данные о том, где находится камера и как она ориентирована. Это особенно важно в робототехнике и в приложениях дополненной реальности, где нужна очень точная навигация. Кроме того, использование IMU делает системы SLAM более надежными и стабильными, особенно в портативных проектах таких как дроны или роботы.

Таким образом, IMU в Intel RealSense T265 обеспечивает надежные данные о движении, которые критически важны для точного определения положения и картографирования окружающей среды, делая его неотъемлемой частью современных систем навигации и отслеживания.

## 3.2 Visual Odometry

Visual Odometry (VO) – технология позволяющая понять, как перемещается устройство, анализируя кадры, снятые камерой [11]. Другими словами, система анализирует фото, чтобы выявить важные моменты или объекты, например, уникальные точки на снимке, и сравнивает их между разными кадрами. Это помогает системе понять, как изменилось положение устройства в окружающем мире.

### 3.2.1 Алгоритмы отслеживания особенностей

Для реализации технологии VO, используются алгоритмы отслеживания особенностей, такие как FAST, SIFT, SURF, и ORB. Эти алгоритмы –инструменты, которые могут быстро находить и следить за определёнными точками или объектами на изображении от одного кадра к другому. Они помогают системе VO точно определить, куда и как двигается устройство, обеспечивая необходимые данные для создания точной карты окружающего пространства. Сравнение нахождения ключевых точек алгоритмами отслеживания особенностей изображено на рисунке 4.

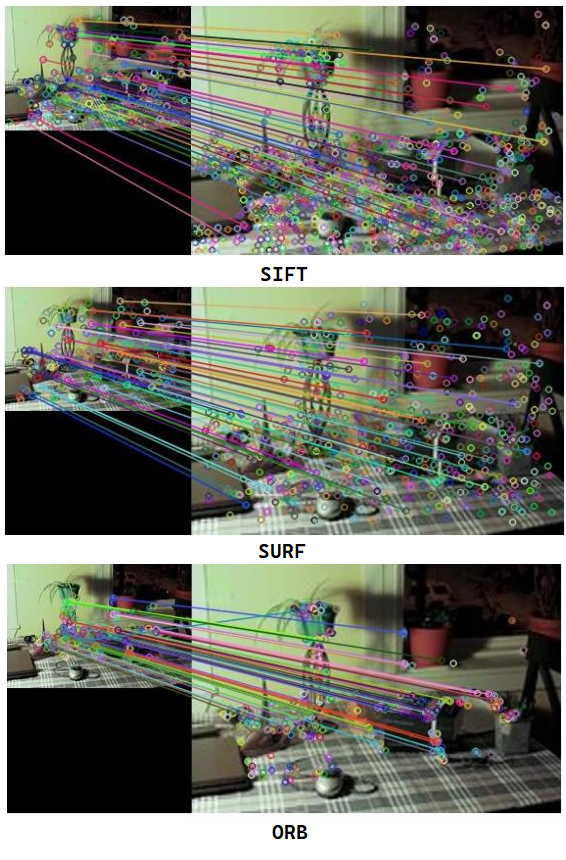


Рисунок 4 – Нахождение ключевых точек алгоритмами SIFT, SURF и ORB

Каждый из этих алгоритмов имеет свои особенности. Например, FAST отлично работает в реальном времени, быстро находя углы и выделяя контрастные участки, но может быть чувствителен к шуму и изменениям освещения. Преимущество SIFT заключается в том, что он может найти особенности, стойкие к изменению масштаба и повороту, делая его отличным выбором для сопоставления кадров между собой. SURF работает похоже на SIFT, но быстрее, благодаря упрощенным вычислениям, а ORB сочетает в себе быстроту и устойчивость, делая его хорошим выбором для систем с ограниченными ресурсами.

Данные алгоритмы необходимы для успешной реализации систем VO и SLAM, поскольку они позволяют точно анализировать движение и ориентацию устройства в пространстве, делая навигацию более надежной и точной, что является незаменимым для автономных транспортных средств, робототехники и приложений дополненной реальности.

### 3.2.2 Вклад Visual Odometry в SLAM

В контексте SLAM технология Visual Odometry является неотъемлемой частью процесса построения 3D карт. Этот метод использует камеры для анализа серии изображений, чтобы понять, как перемещается устройство в пространстве. Это особенно полезно в местах, где GPS-сигналы недоступны или неэффективны, например, внутри зданий или в густонаселённых городских районах. С помощью VO системы могут очень точно отслеживать движение даже в условиях, когда окружающая ситуация быстро меняется.

Тем не менее, у VO есть недостатки, включая зависимость от качества и освещённости изображений. К тому же, чем дольше система использует VO без коррекции, тем больше ошибок может накопиться в измерениях.

VO выполняет критически важную функцию, предоставляя начальные данные о перемещении, которые затем используются для создания и обновления карты местности. Полученная информация дополняется данными от других датчиков, например, инерциальными измерительными устройствами (IMU), что существенно повышает точность и надежность системы локализации и картографирования. Один из ярких примеров эффективности VO – его использование в камере Intel RealSense T265. Здесь VO обеспечивает точное отслеживание изменений положения устройства, анализируя данные с двух рыбьеглазных камер, что позволяет создавать высокодетализированные и точные карты местности.

Подводя итоги, Visual Odometry представляет собой мощный инструмент для точного и надежного определения положения устройства, особенно в условиях, где традиционные средства навигации ограничены или вовсе недоступны. Ее способность анализировать визуальные данные и точно вычислять движение делает VO незаменимой в современных системах SLAM. Алгоритмы отслеживания особенностей, включенные в процесс VO, улучшают способность системы к точному и эффективному картографированию и навигации, повышая ее ценность в различных приложениях, от автономных транспортных средств до робототехники и дополненной реальности.

## 3.3 SLAM

В рамках научно-исследовательской работы, особое внимание уделяется алгоритмам Simultaneous Localization And Mapping (SLAM). Эти алгоритмы обеспечивают двойную функцию — локализацию устройства в пространстве и одновременное построение карты окружающей среды [12]. Научный интерес к SLAM обусловлен необходимостью создания надежных и точных систем навигации для роботов, беспилотных автомобилей и приложений дополненной реальности, способных адаптироваться к динамично изменяющимся условиям окружающей среды.

Суть технологии SLAM заключается в решении задачи оценки состояния системы, которая включает в себя два основных аспекта: оценку положения устройства в пространстве и создание карты окружающей среды. Эти задачи решаются одновременно, что отличает SLAM от обычных методов навигации, где необходима предварительно загруженная карта местности или внешние сигналы для определения местоположения, например, сигналы GPS.

Основная сложность, с которой можно столкнуться при работе с SLAM, заключается в необходимости эффективного слияния данных с множества датчиков, таких как камеры, лидары, радары и инерциальные измерительные устройства (IMU).

Методология исследования в области SLAM включает в себя разработку алгоритмов для эффективного распознавания и отслеживания особенностей окружающей среды, оценки движения устройства между последовательными измерениями и оптимизации полученной карты для минимизации ошибок. Современные подходы к SLAM, такие как фильтр Калмана, частицы фильтра и графические методы, предоставляют различные стратегии для решения этих задач.

Фильтр Калмана используется для слияния шумных данных измерений и предсказания состояния системы на основе линейных моделей. Однако, в условиях нелинейности, которые часто встречаются в реальных сценариях, применяются расширенный фильтр Калмана (EKF) и нелинейный фильтр Калмана, такой как Unscented Kalman Filter (UKF), предлагающие способы работы с нелинейными моделями. Визуализация фильтра Калмана на графике изображена на рисунке 5.

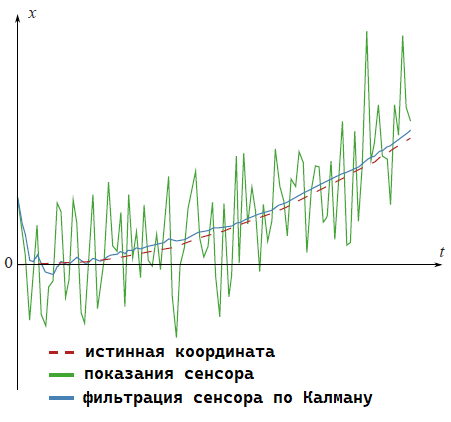


Рисунок 5 – Пример работы фильтра Калмана

Фильтры частиц предоставляют альтернативный подход, основанный на методе Монте-Карло, который эффективен в условиях высокой неопределенности и может обрабатывать многомерные пространства состояний. Этот метод особенно хорошо работает в ситуациях, где пространство состояний системы сложно и многомерно, то есть, когда у нас есть много переменных для учета. Используя метод Монте-Карло — технику, основанную на случайности и вероятностях, фильтры частиц создают множество виртуальных сценариев (называемых частицами), каждый из которых представляет возможное состояние системы. Эти частицы распределяются таким образом, чтобы охватить все возможные исходы, основываясь на текущих данных и измерениях. На рисунке 6 изображён процесс адаптивной локализации Монте-Карло.

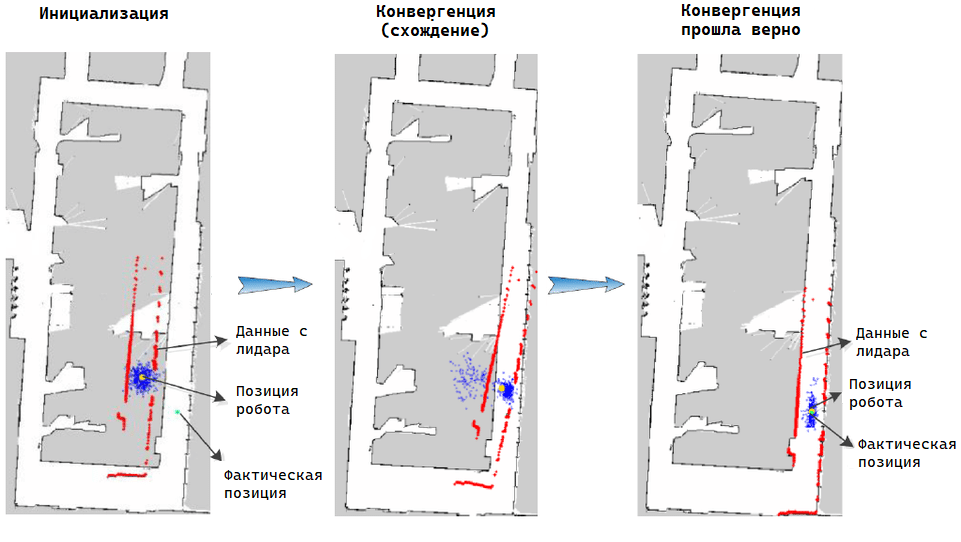


Рисунок 6 – Процесс адаптивной локализации Монте-Карло

Графические методы SLAM, в свою очередь, представляют собой мощный инструмент для оптимизации карт и траекторий, используя графы для представления взаимосвязей между различными элементами карты и измерениями. Эти методы позволяют эффективно решать задачу SLAM путем минимизации ошибок в соответствии с моделью наблюдений и движения.

Подводя итоги, технология SLAM выступает ключевым элементом в развитии автономных систем, обеспечивая их способность к самостоятельной навигации и картографированию в среде, где использование традиционных методов навигации затруднено или невозможно. Прогресс в области SLAM открывает двери для повышения эффективности и точности в робототехнике, автономных транспортных средствах и приложениях дополненной реальности, делая эти системы более независимыми и функциональными.

# Реализация трекинга и навигации

После ознакомления с основами трекинга и навигации, нам необходимо реализовать их на практике. В данном разделе описан процесс работы с камерой Intel RealSense T265, от установки окружения до успешной реализации трекинга и навигации.

## 4.1 Установка окружения

Чтобы осуществить трекинг и навигацию, необходимо установить VMware Workstation Player. Для использования VMware Workstation Player на Windows, требуется активировать функцию виртуализации, которую можно включить так:

1. Войдём в BIOS, нажав клавишу F2 во время загрузки компьютера (на примере Aorus Gigabyte).
2. В меню M.I.T. переходим в раздел Advanced Frequency Settings.
3. Затем выбираем Advanced CPU Settings.
4. Устанавливаем режим SVM в положение Enabled.
5. Сохраняем изменения и перезагружаем систему.

После этого необходимо создать виртуальную машину с операционной системой Ubuntu 22.04 в VMware Workstation Player. Конечные спецификации созданной виртуальной машины представлены на рисунке 7.

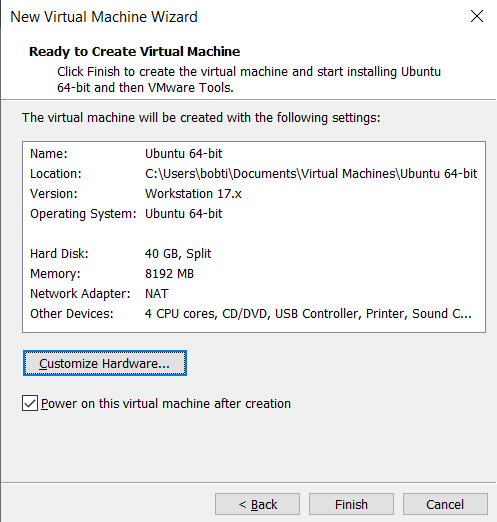


Рисунок 7 – Характеристики виртуальной машины

Установим необходимое окружение для реализации трекинга и навигации (инструкция предоставлена [по ссылке](https://github.com/FroLaCocytus/Intel-RealSense-T265)):

* Установим Cmake и компилятор C++ для сборки проекта.
* Установим Intel RealSense SDK 2.0 для возможности работать с камерой T265.
* Настроим правила udev для получения доступа к нашей камере.
* Установим библиотеку pyrealsense2 для возможности писать свои скрипты для камеры.

## 4.2 Изучение SDK 2.0

После того как мы выполнили вышеописанные этапы, необходимо проверить работоспособность нашей камеры в графическом интерфейсе SDK 2.0. На рисунке 8 изображены показания камеры в 2D плоскости.

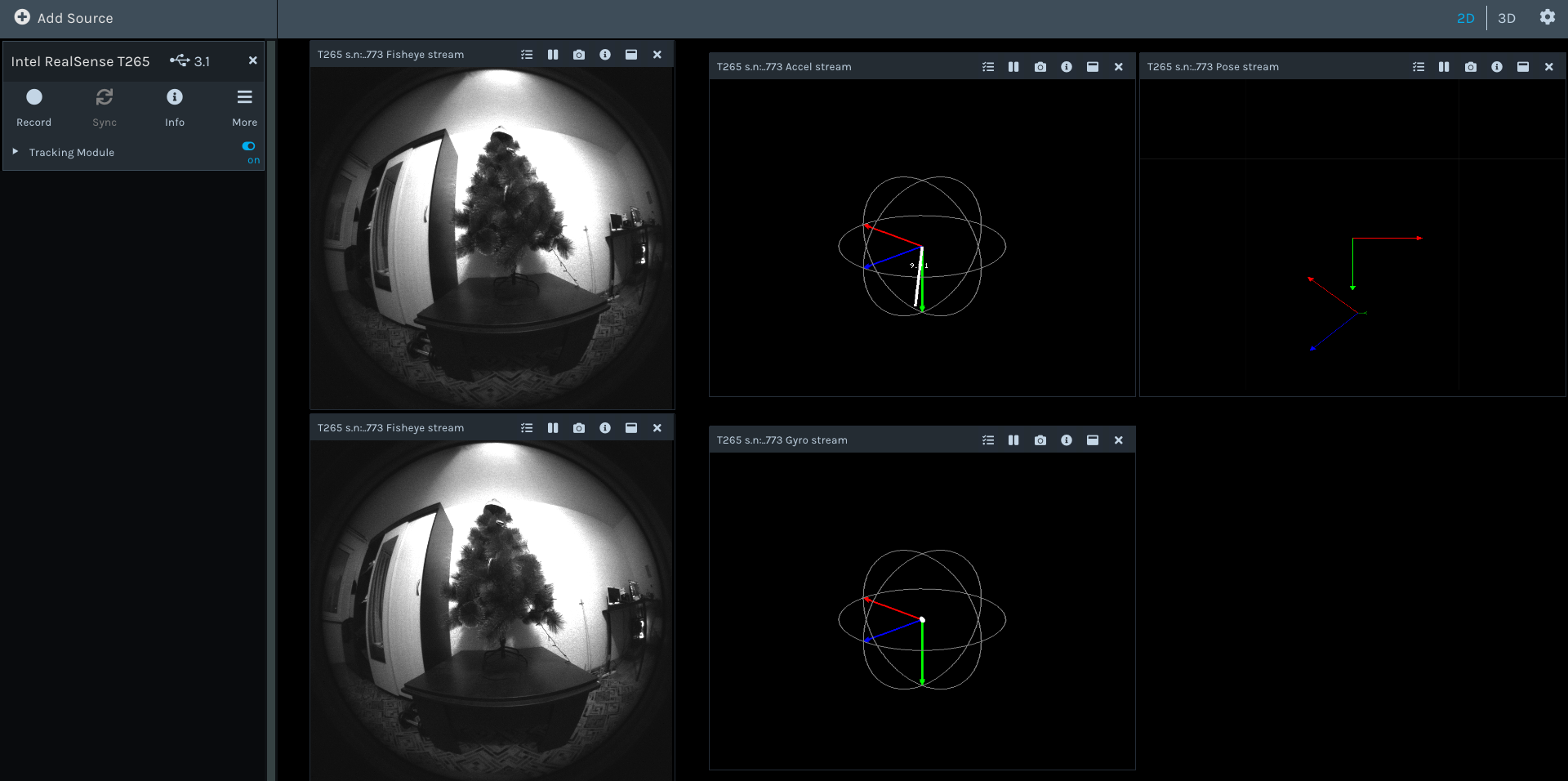


Рисунок 8 – Показания камеры в 2D плоскости

Мы видим данные с двух рыбьеглазных камер, показание с гироскопа, а также позиционирование и ускорение с акселерометра.

Теперь необходимо посмотреть работает ли трекинг у камеры. Для этого откроем 3D плоскость и пройдёмся по помещению. Показания камеры в 3D плоскости изображены на рисунке 9.

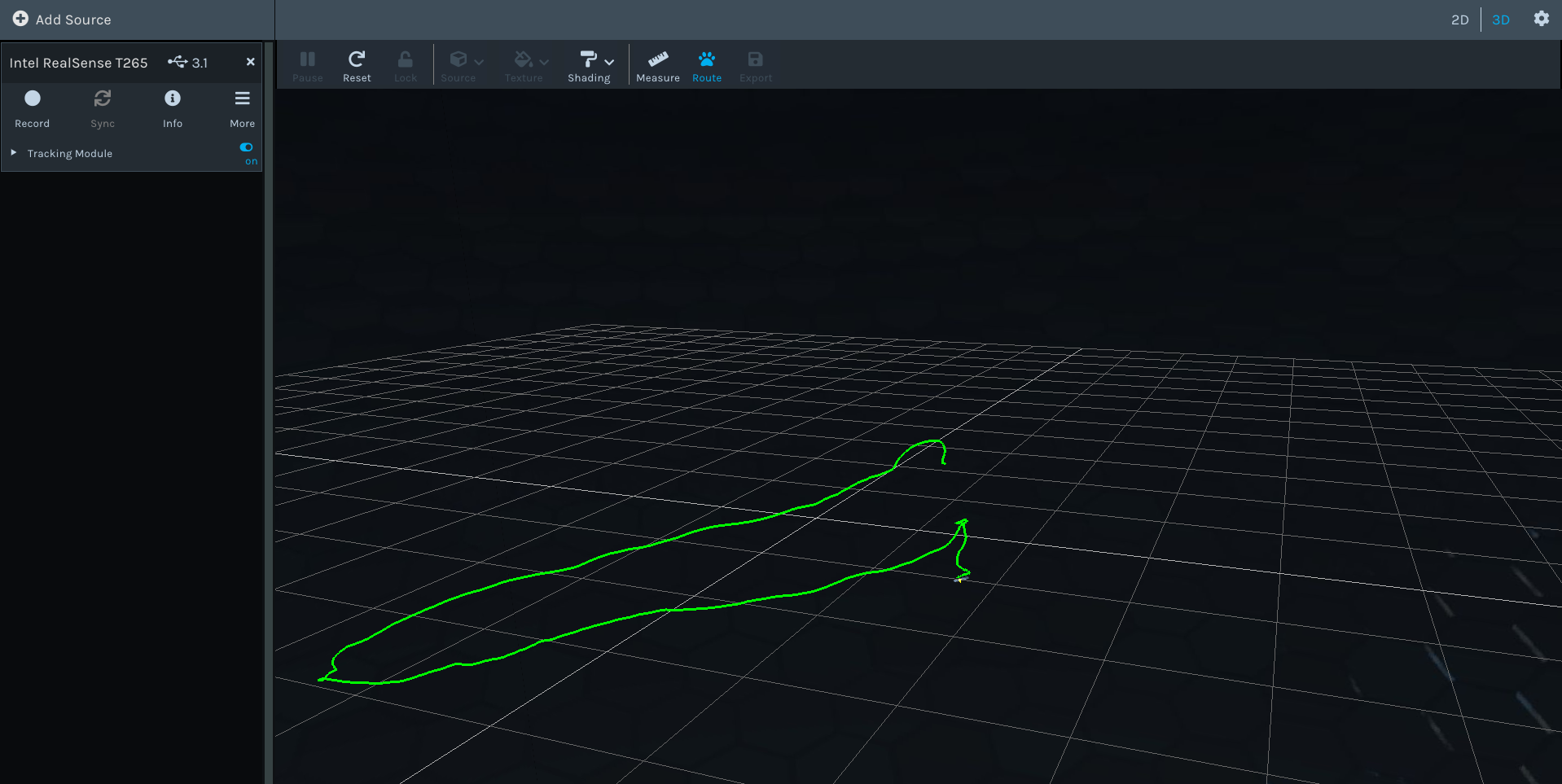


Рисунок 9 – Показания камеры в 3D плоскости

На рисунке видно, что камера успешно выполняет трекинг. Зелёная полоска обозначает что при перемещении камеры не было особых помех. В случае, если мы попытаемся пройтись по тёмному помещению, полоска может стать жёлтой или красной, в зависимости от силы зашумления. Это связано с тем, что наши рыбьеглазные камеры передают недостаточно полноценные данные об окружающем пространстве, и, соответственно, трекинг камеры ухудшается.

## 4.3 Написание скриптов на pyrealsense2

После того как мы убедились в том, что камера выполняет свой функционал, напишем два скрипта на Python чтобы научится работать с данными приходящими с камеры.

Для начала научимся считывать данные с акселерометра и гироскопа. Код скрипта находится [по ссылке](https://github.com/FroLaCocytus/Intel-RealSense-T265/blob/main/t265_pose_tracking.py). На рисунке 10 изображён вывод скрипта с получением данных акселерометра и гироскопа.

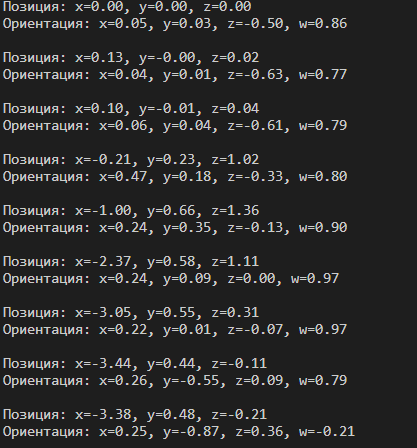


Рисунок 10 – Вывод скрипта с получением данных акселерометра и гироскопа

На рисунке выше мы успешно считали данные с акселерометра и гироскопа. Мы получаем информацию о позиции и ориентации. У позиции мы видим положение камеры в пространстве относительно осей x, y и z. В ориентации мы получаем кватернион, где x, y, z описывают ось вращения в трехмерном пространстве, а w определяет угол вращения вокруг оси.

Теперь напишем более сложный скрипт для трекинга камеры. Мы будем использовать библиотеку matplotlib для визуального отображения маршрута камеры. Код скрипта находится [по ссылке](https://github.com/FroLaCocytus/Intel-RealSense-T265/blob/main/t265_3d_visualization.py). Визуализация трекинга камеры изображена на рисунке 11.

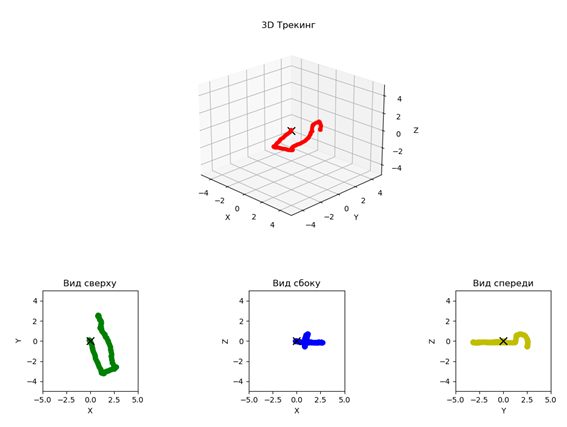


Рисунок 11 – Визуализация трекинга камеры

На рисунке выше, мы перемещались из одной комнаты в другую находящуюся за стеной. В конце мы подняли камеру вверх и затем опустили. Камера успешно считала это и отобразила на графиках. Крестиком обозначена изначальная позиция камеры. Для более расширенного понимания добавим вручную стенку на график «3D Трекинг» и «Вид сверху». На рисунке 12 добавлены пояснения к трекингу камеры.

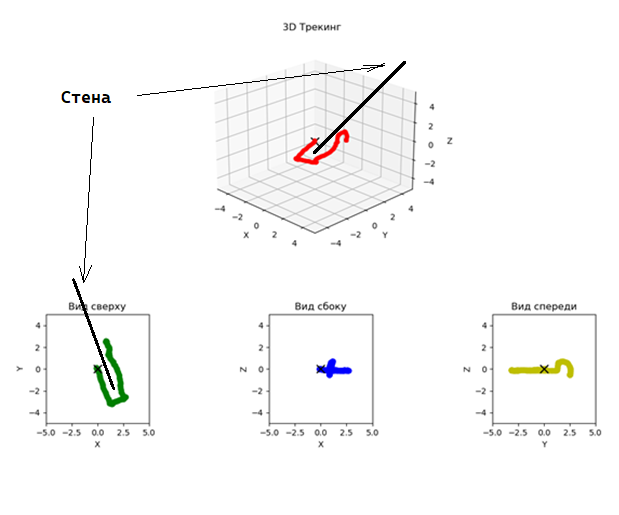


Рисунок 12 – Пояснение к трекингу камеры

На рисунке видно, что камера обогнула стену. Таким образом, мы убедились, что камера действительно отобразила свой маршрут верно.

Подводя итоги, мы научились основам работы с камерой T265. Смогли успешно реализовать навигацию и трекинг камеры. Научились работать с данными с камеры и написали два скрипта один из которых визуализировал маршрут камеры в помещении.

# Заключение

В ходе выполнения настоящей научно-исследовательской работы был проведён анализ существующих научных публикаций по вопросам трекинга и навигации и смежных им тем. Помимо этого, были детально рассмотрены основополагающие алгоритмы и методы, используемые камерой IntelRealsense T265, такие как: SLAM, Visual Odometry и IMU Fusion. Изучение технических характеристик, принципов работы и возможностей применения камеры Intel RealSense Tracking Camera T265 позволило подтвердить её важность для решения задач трекинга и навигации в трехмерном пространстве.

Важную часть работы составила практическая реализация трекинга и навигации с использованием этой камеры. Был описан процесс установки необходимого ПО, что создаёт фундамент для будущих исследований и разработок в данной области. Результаты экспериментов показали, что камера T265 эффективно справляется с задачами навигации, что соответствует предварительным ожиданиям относительно её производительности и точности.

Таким образом, камера Intel RealSense T265 демонстрирует значительный потенциал для разработки систем трекинга и навигации, предоставляя точные и надёжные решения, пригодные для множества приложений. На основе полученных результатов планируется продолжение исследований в данной области. Особое внимание будет уделено разработке интегрированной системы, которая объединяет возможности двух камер: Intel RealSense T265 для трекинга и навигации и Intel RealSense D435 для глубинного восприятия. Это сочетание камер позволит нам сделать усовершенствованную систему, которая будет лучше справляться в трудных ситуациях, делая устройства более самостоятельными и улучшая их способность точно определять своё местоположение в пространстве.

# Список использованных источников

1. Hausamann P., Sinnott C., Daumer M., Macneilage P. Evaluation of the Intel RealSense T265 for tracking natural human head motion // Scientific Reports – 2021.
2. Thomas G., Gade R., Moeslund T., Carr P. Computer Vision for Sports: Current Applications and Research Topics // Computer Vision and Image Understanding – 2017.
3. Ren C., Prisacariu V., Kähler O., Reid I., Murray D. Real-Time Tracking of Single and Multiple Objects fromDepth-Colour Imagery Using 3D Signed Distance Functions // International Journal of Computer Vision – 2017. P.80 – 96.
4. Pesenti M., Invernizzi G., Mazzella J., Bocciolone M., Pedrocchi A., Gandolla M. IMU-based human activity recognition and payload classification for low-back exoskeletons // Scientific Reports – 2023.
5. Aqel M., Marhaban M., Saripan M., Ismail N. Review of visual odometry: types, approaches, challenges, and applications // SpringerPlus – 2016.
6. Zhao X., Agrawal H., Batra D., Schwing A. The Surprising Effectiveness of Visual Odometry Techniques for Embodied PointGoal Navigation // International Conference on Computer Vision – 2021.
7. Bayer J., Faigl J. On Autonomous Spatial Exploration with Small Hexapod Walking Robot using Tracking Camera Intel RealSense T265 // European Conference on Mobile Robots – 2019.
8. Datasheet Intel RealSense T265 [Сайт]. – URL: https://dev.intelrealsense.com/docs/tracking-camera-t265-datasheet
9. Акселерометр [Сайт]. – URL: <https://en.wikipedia.org/wiki/Accelerometer> – Текст: электронный.
10. Гироскоп [Сайт]. – URL: <https://en.wikipedia.org/wiki/Gyroscope> – Текст: электронный.
11. Visual Odometry [Сайт]. – URL: <https://www.ifi.uzh.ch/en/rpg/research/research_vo.html> – Текст: электронный.
12. SLAM [Сайт]. – URL: <https://ru.wikipedia.org/wiki/SLAM_(метод)> – Текст: электронный.