

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ УЧРЕЖДЕНИЕ
«ФЕДЕРАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЦЕНТР ИНСТИТУТ
ПРИКЛАДНОЙ МАТЕМАТИКИ ИМ. М.В. КЕЛДЫША РОССИЙСКОЙ
АКАДЕМИИ НАУК»

На правах рукописи

Подопросветов Алексей Валерьевич

ОПРЕДЕЛЕНИЕ ПАРАМЕТРОВ ДВИЖЕНИЙ ЧЕЛОВЕКА

Специальность 1.1.7 —
«Теоретическая механика, динамика машин»

Диссертация на соискание учёной степени
кандидата физико-математических наук

Научный руководитель:
канд. физ.-мат. наук, доцент
Орлов Игорь Александрович

Москва — 2023

Оглавление

	Стр.
Введение	5
Глава 1. Разработка системы сбора кинематических данных на основе видеоданных	31
1.1 Цели и задачи	31
1.2 Описание эксперимента по сбору видеоданных	31
1.3 Восстановление трёхмерной скелетной модели человека по видеоданным	35
1.3.1 Обработка кадров OpenPose	35
1.3.2 Постановка задачи	38
1.3.3 Определение координат точек	41
1.3.4 Использование геометрической информации о скелете для улучшения распознавания	42
1.3.5 Контроль восстановления скелета	44
1.3.6 Фильтрация полученных данных	46
1.3.7 Интерполяция	47
1.4 Результаты	48
Глава 2. Анализ последовательностей движений человека на основе видео данных	49
2.1 Классификация последовательностей движений человека	50
2.1.1 Формирование датасета	51
2.1.2 Результаты классификации	53
2.2 Прогнозирование следующей последовательности движений	54
2.2.1 Набор обучающих примеров	54
2.2.2 Результаты	56
2.3 Прогнозирование начала следующей последовательности движений	57
2.3.1 Предобработка обучающих наборов данных	58
2.3.2 Результаты прогнозирования	61

2.3.3	Заключение	63
Глава 3. Система по сбору кинематических данных на основе датчиков MEMS 64		
3.1	Введение	64
3.2	Описание биомеханической модели тела человека	66
3.2.1	Упрощение модели тела человека	69
3.2.2	Механическая модель тела человека	72
3.3	Эксперимент по сбору данных с видео и MEMS	76
3.3.1	Описание MEMS датчиков BNO055	77
3.3.2	Получение ориентаций из данных датчика BNO055	80
3.3.3	Сбор и предобработка данных с MEMS датчиков	84
3.3.4	Визуализация данных	87
3.4	Вычисление динамических характеристик движения	88
3.4.1	Решения задачи динамики	88
3.4.2	Оценка энергии, затрачиваемой человеком	92
3.5	Алгоритм калибровки микромеханических датчиков к сегменту	103
3.5.1	Расположение датчиков на теле человека	105
3.5.2	Этапы калибровки	106
3.6	Практическое применение системы сбора данных	108
3.6.1	Система визуализации данных	109
3.6.2	Предобработка кинематических данных в формате R3DG	111
3.6.3	Предобработка кинематических данных в формате kinect (.skeleton)	112
3.7	Результаты	115
Глава 4. Анализ последовательностей движений человека на основе совокупности видео данных и данных микроэлектромеханических датчиков 117		
4.1	Классификация последовательностей движений человека	117
4.1.1	Целевые переменные	119
4.1.2	Формирование общего набора данных	120
4.1.3	Получение дополнительных признаков	120
4.1.4	Подготовка обучающей выборки	123
4.1.5	Результаты обучения	124

4.1.6	Вычислительная сложность и оценка времени работы . . .	126
4.1.7	Заключение	128
4.2	Сравнительные характеристики для разных источников данных .	128
4.2.1	Результаты сравнения моделей для разных источников данных	130
4.3	Прогноз вращений и ускорений для элементов модели тела человека	133
4.3.1	Постановка задачи	134
4.3.2	Описание эксперимента	135
4.3.3	Реализация описанного метода	136
4.3.4	Результаты работы алгоритмов прогнозирования	139
	Заключение	141
	Список сокращений и условных обозначений	143
	Словарь терминов	147
	Список литературы	149
	Список рисунков	162
	Список таблиц	164
	Приложение А. Демонстрация результата обработки видео данных эксперимента по сбору последовательностей движений	165

Введение

С развитием технологий растет и число возможных применений систем биомеханической оценки деятельности человека, в частности их используют для взаимодействия человека с технологическим оборудованием и роботами, для анализа производительности труда и контроля безопасности рабочих на предприятиях, для определения профзаболеваний, для реабилитации в спортивной медицине, для формирования реалистичных движений в кино, компьютерных играх и пр. Принципиальная особенность технологического прогресса состоит в необходимости изобретать новые способы обработки и анализа данных, а также методы их адаптации под определенные нужды.

Н. А. Бернштейн (1896-1966) по праву может считаться создателем теоретической основы современной биомеханики — учения о двигательной активности людей и животных. Он фиксировал рабочие операции, такие как стук молотком, бытовые движения (еда ложкой), музицирование (различный темп и интенсивность звука), а также различные виды перемещения, включая ходьбу неврологических больных, стариков, младенцев и детей старшего возраста, сравнивая их движения к перемещению у здоровых взрослых. Экспериментальные исследования Бернштейна были тесно связаны с появившимися в конце девятнадцатого века новыми способами анализа и регистрации движения, такими как техника хронофотографии, циклометрию. Работа Бернштейна открыла новую дисциплину под названием кинезиология, изучающую структуру и механизмы движения. Он предложил теорию «уровней контроля над построением действия», которая позже нашла отражение в уровнях контроля в когнитивной психологии. Бернштейн также ввел термин биомеханика, изучение движения посредством применения механических принципов.

Также ученым перешедшим из области математики в область биологии был И. М. Гельфанд. Он едва ли не уникальный пример самообразования, завершившегося блестящей научной карьерой. Без формального среднего и высшего образования он стал аспирантом А. Н. Колмогорова и выдающимся математиком своего времени. Автор фундаментальных результатов в области теории обобщённых функций, занимался дифференциальными уравнениями, теорией топологических линейных пространств, обратными задачами спектраль-

ного анализа, квантовой механикой, динамическими системами, имеющими существенное значение для теоретической физики. Гельфанд работал над применением математических методов описания поведения сложных систем к изучению регуляции механизмов управления локомоцией у млекопитающих. Вместе с соавторами им были предложены принцип наименьшего действия и представления о синергиях в управлении деятельностью элементов сложных биологических систем [1].

Другой пионер современного этапа исследования движений Гурфинкель В.С. внес заметный вклад в развитие метода стабилотрии. В 1970 году за участие в создании протеза предплечья с биоэлектрическим управлением удостоен Государственной премии СССР. Основные области научных интересов физиология, биофизика движений и робототехника. Часто биомеханику движений человека разделяют на три аспекта: теоретико-механический, медицинский и биофизический. В представленной работе будут рассматриваться способы фиксации физической деятельности человека и строиться упрощенная модель сегментов тела человека.

Существенный вклад в изучение проблемы восстановления траекторий движения инерциальных сенсоров был внесен [2]. Хотя работа не относится к восстановлению кинематики движений модели человека, однако предлагает метод многопозиционной калибровки для инерциальных датчиков, не требующий специального ориентированного крепления, который был адаптирован для компенсации большинства детерминированных ошибок датчика, включая важный масштабный коэффициент и ошибки неортогональности гироскопов. Другой немаловажный вклад был представлен двумя группами [3], [4] и [5], [6] где используется автокалибровка ошибок интегрирования, связанных с динамикой их движения при условии, что датчики закреплены на подвижной системе. Более совершенный метод предложен [7] и [8], который заключается в одновременном выполнении трехмерной оценки положения и калибровки инерциальных датчиков в реальном времени на основе методов оптимизации в неизвестных условиях без какого-либо внешнего оборудования. При всей общности, данный метод создан для инерциальной навигации и не учитывает возможные кинематические связи в системах датчиков в отличие от [6]. Однако в его работе не учитываются в явном виде погрешности, вносимые расположением датчиков относительно частей тела человека, при этом не явно частично решается

задача их оптимизации. Восстановление траекторий движения инерциальных сенсоров, предложенное далее автором данной работы, кардинально отличается свободой расположения датчиков на частях тела человека, что обобщает предложенные ранее подходы, но требует алгоритма предварительной калибровки. Построенный алгоритм решает задачу оптимизации относительного положения датчиков на сегментах тела человека, и включает описанные выше наработки по калибровке и учету погрешностей статических и детерминированных ошибок инерциальных датчиков.

Основные задачи данной диссертационной работы, в целом посвященной анализу последовательностей движений человека в условиях работы на промышленном предприятии, могут быть разбиты на несколько аспектов.

В первую очередь – это разработка инструмента промышленного образца для сбора, хранения и использования кинематических скелетных данных (о движении основных частей тела и суставов их соединяющих в трехмерном евклидовом пространстве). Описание подробных алгоритмов исключения статических, динамических и детерминированных ошибок при восстановлении положения и ориентации инерциальных систем широко представлены в разных областях, в частности, для инерциальной навигации.

Понимание статических, динамических и детерминированных ошибок для IMU датчиков важно при их использовании для измерений и навигации. Статические ошибки: Статические ошибки (или нулевые ошибки) возникают, когда IMU датчик выдает ненулевые значения, даже если объект, на котором установлен датчик, находится в покое (не двигается и не вращается). Эти ошибки могут возникать из-за дрейфа нулевой точки (*offset*) датчика или из-за несовершенства калибровки. Статические ошибки могут быть скорректированы путем калибровки и компенсации нулевых значений. Динамические ошибки: Динамические ошибки (или динамический дрейф) возникают, когда IMU датчик показывает изменения ускорения или угловой скорости, даже если объект находится в покое. Эти ошибки могут быть вызваны различными факторами, такими как вибрации, температурные изменения, электромагнитные воздействия и другие. Для уменьшения динамических ошибок обычно используются фильтры и алгоритмы компенсации, а также более точные датчики. Детерминированные ошибки: Детерминированные ошибки – это ошибки, которые можно предсказать и объяснить. Они имеют определенную причину и могут быть воспроизведены

и учтены в алгоритмах коррекции и фильтрации данных. Примером детерминированных ошибок может быть известное электромагнитное воздействие на датчик, которое вызывает изменения в его показаниях. Детерминированные ошибки могут быть устранены путем моделирования и коррекции с учетом известных причин.

Алгоритмов калибровки инерциальных систем по отношению к устройствам, где они установлены, значительно слабее представлены в литературе и еще меньше алгоритмов калибровки инерциальных систем, одеваемых на человека. Существует классификация подходов для калибровки: технический, статический и функциональный, предложенная авторами работы [9]. Они описывают проблему недостатка методических работ, позволяющих воспроизвести описываемые способы калибровки, а также проблему оценки точности существующих алгоритмов, так как каждый автор предлагает свой протокол ее оценки. Однако для оценки предлагаемых методов для конкретного приложения необходимо провести их апробацию и сравнение с использованием одного протокола. Автором данной работы предлагается алгоритм, который объединяет все три подхода из классификации [9], и может оцениваться по любому протоколу.

Во-вторых, в работе предложена система методов анализа данных последовательностей движений человека с использованием робастных моделей машинного обучения не требующих больших обучающих выборок, что позволяет обобщить полученные результаты для использования при анализе деятельности рабочих на промышленных предприятиях. Методы машинного обучения для детектирования конкретных последовательностей движений, которые дальше будут называться действиями и предсказания последующего действия широко представлены в литературе (например [10]). Работ с предсказанием момента начала последующей последовательности движений значительно меньше, но они тоже есть [11], Из-за того, что для каждого рабочего места существует свой набор таких последовательностей, известные работы посвящены или концептуальному описанию широких классов действий, или хорошо формализуемым методам исследования в узких классах. Одним из результатов данной работы является расширение классов действий, которые могут быть формализованы без разработки новых методов анализа данных последовательностей движений человека.

В-третьих, прогнозирование кинематических характеристик сегментов тела человека с целью создания альтернативного движения системой управления механического устройства с заданным параметром оптимизации имеет множество аналогов в литературе, особенно в области разрешения прямых и обратных задач кинематики и динамики для манипуляторов. В данной исследовательской работе демонстрируется пример нового подхода к усовершенствованию традиционного метода разработки устойчивых роботизированных систем. Предложенный метод основывается на интеграции методов искусственных нейронных сетей, в математические модели, с целью сглаживания ограничений, связанных с созданием этих моделей. Простая структура предложенного подхода дает возможность применение при разработке широкого спектра робототехнических систем.

При создании мышечно-скелетной модели для симуляции и изучения движения человека важно определить компромисс между сложностью модели и скоростью симуляции. Модели с приводом по моменту использовались для решения задач оптимального управления [12]; хотя такие модели предоставляют ценные сведения и обладают высокой вычислительной скоростью, они ограничены в своей биологической достоверности. С другой стороны, конечно-элементные симуляции мышц хорошо представляют сложную геометрию мышц, но они требуют больших вычислительных затрат [13]. Между этими крайностями находятся модели типа Хилла и модели с трехмерной скелетной геометрией. Чаще других такие модели строят для нижней части тела человека Delp et al. [14], Klein Horsman et al. [15], Arnold et al. [16] и Carbone et al. [17], и включают в системы моделирования такие как OpenSim [18], AnyBody [19].

Еще одним важным аспектом является выбор соответствующих экспериментальных данных для определения параметров модели. Модель, созданная Delp et al. [14], широко используется для создания симуляций ходьбы и других движений нижних конечностей, но параметры мышц и сухожилий (например, длина, сила, максимальная скорость) не всегда определены точно и могут значительно варьировать у различных пациентов или испытуемых. Однако мускулоскелетные модели становятся все более точными и реалистичными, что позволяет вернее смоделировать и изучить движение человека. Возможности применения этой технологии включают спортивные науки, реабилитацию и ме-

дицину, а также создание более точных и реалистичных виртуальных человекоподобных моделей.

Некоторые модели скелетно-мышечной системы, такие как разработанные Arnold et al. [16], стали широко используемыми, но они имеют свои недостатки для использования в симуляциях, основанных на действии мышц. Например, для точного представления геометрии мышц эта модель использует поверхности эллипсоидальной обертки, что делает вычисление путей мышц вычислительно затратным. Кроме того, оптимальные значения сил мышц в таких моделях могут быть не совсем точными, так как они основаны на данных, полученных от пожилых людей с существенной атрофией мышц, что может недостаточно точно отражать распределение сил в мышцах молодых здоровых людей.

Однако существуют более современные подходы, которые используют более актуальные данные о массе и объеме мышц у молодых здоровых испытуемых, полученные с помощью магнитно-резонансной томографии (МРТ). Такие данные позволяют создать более реалистичные модели мускулоскелетной системы, учитывая разнообразие параметров у различных людей. Это обеспечивает более точные результаты и дает надежность использования модели в различных приложениях, таких как спортивные науки, реабилитация и медицина.

Другой проблемой является механическое правдоподобие воспроизведение человеческого движения. В контексте человеко-машинных интерфейсов, для этого используются соматосенсорная информация, в частности для управления устройствами или роботами и создания более естественного взаимодействия между человеком и машиной. Соматосенсорные рецепторы расположены по всему телу человека и отвечают за восприятие различных физических и тактильных аспектов окружающего мира, таких как прикосновения, давление, температура, болевые ощущения и положение различных частей тела, они передают информацию в мозг, где она обрабатывается и интерпретируется. Эта информация играет важную роль в координации движений, оценке окружающей среды и обеспечении безопасности и защиты организма. Соматосенсорная информация также позволяет создавать более реалистичные модели людей [20; 21], предоставляя более подробное внутреннее состояние. Ряд исследований в области анализа и моделирования движений скелетно-мышечной модели человека проводились в области спортивных наук и медицины [22; 23]. Их основная цель

– использование соматосенсорных данных человека для улучшения движений спортсменов и контроля за восстановлением людей проходящих реабилитацию.

Большинство из них применяют либо упрощенную модель всего тела, либо подробную модель, но небольшой части тела, из-за высокой вычислительной потребности при использовании подробных моделей человеческого тела. Вычислениям всей соматосенсорной информации посвящены исследования [24–26], хотя эффективности вычислительных алгоритмов не уделялось большого внимания.

В данной работе рассматривается многозвенная система без ограничения общности на основе модели предложенной в публикации [27], состоящая из твердотельного скелета и мышечного скелета. Для модели построен закон движения на основе которого решается задача обратной динамики на вход подаются данные о кинематике движения на основе видео камер или инерциальных датчиков, в результате вычисляются силы и моменты в моделях мышц.

В настоящее время одним из самых распространенных и глубоко изученных решений по сбору кинематических данных является оптоэлектронная система (*motion capture*). Данная технология основана на инфракрасных камерах, которые отслеживают трехмерное положение маркеров, размещенных на объекте. В свою очередь, оптические системы делятся на два типа: пассивные и активные.

К пассивным относятся системы с отражающими маркерами. К активным относятся системы с маркерами-излучателями, в которые включены контроллеры управления, необходимые для синхронизации и идентификации.

Также бывают схожие системы, где маркерами являются магниты, а камерами – ресиверы. В данном случае система рассчитывает позиции маркеров по искажениям магнитного потока.

Высокая точность измерения таких технологий подтверждается данными, полученными в работе [28] (положение 0,1 мм). Собранные данные возможно использовать для анализа кинематических характеристик человека посредством специализированных программных продуктов, для медицинских и спортивных целей. Однако, существует ряд серьезных ограничений, затрудняющих использование таких систем. К ограничениям можно отнести длительное время калибровки камер или ограниченную область измерения, что не позволяет использовать системы в заранее не подготовленных условиях.

Другим распространенным решением задачи восстановления движений является анализ видеоряда посредством машинного обучения. Ярким примером подобного алгоритма является система OpenPose [29]. К достоинствам данной системы относится её возможность поиска неограниченного числа людей на кадре путем обнаружения ключевых точек на человеке в реальном времени, а также умение распознавать не только человека целиком, но и его части, если человек не поместился в кадр полностью (например, кисть руки или лицо). Однако, разработанная система 3D-распознавания работает только в лабораторных условиях, поскольку система строится по нескольким кадрам с нескольких, заранее известных, ракурсов и работает только на преднастроенных кадрах. К сожалению данные полученные с этой системы остаются непригодными для анализа кинематики полученного движения.

В другой системе 3DHP [30] представлен подход к восстановлению 3D-позы человека по одной камере. Данный подход основан на алгоритмах и архитектуре, описанных в статье [30]. Однако, с тех пор произведены значительные изменения: алгоритм переписали на *python 3.6* и *pytorch 0.4.1*, изменили структуру встроенных подсетей. Описанная система основана на оценке трехмерной позы из двумерных фотографий, снятых в реальных условиях, посредством трехмерных геометрических ограничений. При всех достоинствах, остается ряд причин включающий шум и неоднозначности интерпретации скелета человека из-за чего данные не пригодны для исследования кинематики движения.

Третий способ восстановления движений человека осуществляется при помощи набора инерциальных датчиков (*MEMS*). Такой подход требует отдельного рассмотрения. Существует ряд обзоров, например [28], в которых проводится сравнительный анализ алгоритмов, основанных на данном способе, и верификация различных калибровок.

Авторы обзора фокусируются на калибровке «датчиков к сегменту» для верхней части тела человека, а именно, на оценке точности углов запястья, локтя и плечевого сустава. Модели оцениваются по двум критериям: достоверность (близость к системе отсчета) и точность (воспроизводимость). Кроме того, для оценки используется экспериментальный протокол, включающий в себя изменчивость субъекта и оператора. Но ни в обзоре, ни в статьях из ссылок обзора алгоритма реализации данного вида калибровок нет.

Основным разработчиком, полностью реализующим систему на *MEMS* сенсорах, является *XSens* [31; 32] – поставщик 3D-устройств захвата движения и инерциальных датчиков на основе технологии *MEMS*. Ясно что из-за того, что продукт коммерческий открытого кода системы, компания не предоставляет. В множестве работ существует только упоминание о системе синтеза данных ориентации из инерциальных датчиков. Ориентируясь на патент [33], можно сделать вывод, что разработчики не только используют информацию об ориентации для восстановления кинематики, но и интегрируют показания акселерометров для отслеживания всех типов движений, включая прыжки, смещения в пространстве относительно начальной точки и движение по неровной местности – например, движение по лестнице. Авторы утверждают, что данная система при помощи инерционного зондирования позволяет регистрировать движение плеча, которое не ведет себя как жесткий шарнир. На основе их разработок возможно исследовать кинематику движений человека, но открытых данных или подробных публикаций по их исследованиям нет. К отрицательным эффектам добавляется, то что стоимость их продукта значительная и один полный комплект датчиков, предназначенный для исследований стоит более 26 000\$.

Российским ближайшим конкурентом компании *xsens* является компания *Neurocor* [34]. Деятельность компании направлена на медицинское применение наборов *MEMS* – датчиков для сбора информации о патологиях движения человека. *Neurocor* в основном разрабатывает узкоспециализированные методики для медицинского применения и не исследует прикладные задачи для промышленности.

В продолжении восстановления динамики человека, в статье [35] представлен способ совмещающий в себе две обработки изображений с камер: визуальную одометрию и визуальный семантический анализ. За счет того что камера снимает от первого лица используя динамическую систему у авторов удастся оценить активную силу и момент, которые косвенно запечатлены в данных камеры.

В данной работе рассматривается задача восстановления кинематики движений человека в направлении к медицинским и промышленным целям таким как подсчет энергии затрачиваемой человеком на определенный вид работ, оценка эргономичности рабочего пространства и оценки профессиональных заболеваний.

Задача восстановления движений человека с использованием инерциальных датчиков является остро актуальной для практического применения и на данный момент не до конца изученной.

Обращаясь к задаче восстановления трехмерных моделей человека на основе видео данных стоит рассмотреть несколько основных подходов к решениям. Один из подходов основывается на глубоких нейронных сетях: публикация [36] посвящена разработке метода оценки антропометрических параметров человека на основе глубоких нейронных сетей с использованием адверсариальной аугментации данных. Авторы предложили использовать адверсариальную аугментацию данных для улучшения производительности модели оценки антропометрических параметров человека. В рамках этого подхода модель генерирует дополнительные изображения человека, используя случайные параметры исходных изображений, а затем использует их в процессе обучения. Для тестирования метода были использованы две базы данных, содержащие изображения людей в различных позах. Эксперименты показали, что использование адверсариальной аугментации данных значительно улучшает производительность модели по сравнению с традиционными методами обучения. Результаты работы могут быть использованы в таких областях, как производство одежды, фитнес-индустрия и медицинская диагностика. В статье [37] рассматривается метод автоматического перепозирования человеческих моделей, основанный на глубоких нейронных сетях, который позволяет генерировать новые позы, сохраняя при этом естественный внешний вид и физиологические ограничения. Для обучения модели использовалась база данных, содержащая 3D-модели людей в различных позах. Модель осуществляет циклическое преобразование между исходной и целевой позами, учитывая консистентность 3D формы, позы и внешнего вида, что позволяет получать более реалистичные результаты, чем традиционные методы репозиционирования. Эксперименты показали, что предложенный метод эффективен и может быть использован в таких областях, как анимация, робототехника и виртуальная реальность. Другой основной подход представлен в статье [38]. Описывается метод восстановления трехмерного объекта из двухмерных изображений на основе томографического подхода в обработке сигналов. Область, где расположен объект, разбивается на маленькие объемы, которым присваиваются номера. Затем контролируемая область представляется в виде вектора-оригинала, количество компонент которого равняет-

ся количеству элементов разрешения в контролируемой области пространства. Значение каждой компоненты равняется 1, если элемент разрешения с номером этой компоненты лежит в теле объекта, и 0, если не лежит. Главным этапом решения задачи является восстановление вектора-оригинала по его отображению при помощи алгебры логики. В результате, в работе получены логические выражения, которые позволяют восстановить вектор-оригинал и реализовано компьютерное моделирование метода. Преимуществами данного метода являются отсутствие требования к определенной подсветке, выявлению соответственных точек и использованию нескольких изображений. Недостатком является сложность в реализации данного метода. И последний подход к которому относится и разработанный автором метод представлен в исследовании [39]. Описывается алгоритм построения трехмерной модели сцены на основе стереоснимков. Алгоритм включает выявление соответствующих точек, вычисление их трехмерных координат, вычисление фундаментальной матрицы и матриц камер, а также триангуляцию для полученного облака точек объекта. Недостатками метода являются сложность выявления соответствующих точек и точность вычисления фундаментальных матриц автоматизированным методом.

Другой задачей является поиск и определение последовательности движений человека. Основные результаты получены в работах [40], [41]. Работы фокусируются на статистически распространенных действиях человека, но в действительности прикладные задачи требуют другой более конкретной специфики конкретные примеры можно найти в задачах для медицины [42] или в задачах автора данной работы. В условиях производства требуется выполнять все три этапа сбор обучающих примеров, построение и обучение моделей и их апробация, что накладывает существенные ограничения на каждый из этапов в отличие от лабораторных условий, упомянутых в работах выше. Для построения моделей машинного обучения для решения задач определения действий, предсказания действий и предсказания моментов начала действий было проведено три опытных эксперимента с общим участием более двадцати шести добровольцев по сбору данных для обучающих выборок, построены необходимые модели и проведены эксперименты по их апробации, что и является основным отличием предлагаемого подхода.

Несмотря на то что распознавание человеческой деятельности (HAR) является активно разрабатываемой областью исследований уже более двадцати

лет, все еще существуют ключевые аспекты, решение которых может привести к значительному изменению в данной области.

Уровень разработанности на 2013 год в области HAR на основе носимых датчиков приводится в обзоре [40]. Именно эта область интересует авторов, так как существуют более точные способы восстановления движений человека на основе методов захвата движения, но такие подходы не универсальны и реализуемы лишь в лабораторных условиях, имеют редкое применение в медицине и кинематографии. Автор сосредоточен на более доступном методе для широкого применения с целью повышения качества условий труда рабочих на промышленном предприятии. Биомехатронные системы или биомеханические роботы или биомеханика в области робототехники встречаются в работах А.М. Формальского, Ю.Ф.Голубева, В.Е. Павловского и др., Ravi Vaidyanathan, S. Collins, Hugh M. Herr et al..

Следует отметить, что большинство исследований по распознаванию механических действий до 2016 года и работы J. L. Reyes-Ortiz et al. было сосредоточено на основных видах деятельности (ОД), а переходные виды деятельности (ПД) не принимались во внимание из-за их непродолжительности по сравнению с ОД. Однако обоснованность этого предположения зависит от конкретных приложений. В частности, падения являются наиболее распространенным типом несчастных случаев среди пожилых людей, которые обычно происходят во время постуральных переходных процессов, таких как переход от сидячего положения к стоячему [42]. Более того исследования в этой области часто опирались на носимые датчики основанные на инерциальном сенсоре в смартфоне, однако такие сенсоры лишены будущего в промышленном применении так как не стандартизированы. Проблемы инерциальных систем слежения и навигации описаны в работах Ю.В. Болотина.

С 2016 до 2019 г. публикационная активность в области HAR резко возросла и разбилась на разные прикладные направления, такие как медицина, спорт, промышленная безопасность [43] и кино-индустрия. Наиболее значимые результаты, основанные на глубоких нейронных сетях, были описаны в работе [41]. Одновременное использование инерциальных блоков и видеоанализа представлены в работах П.А. Кручинина для определение взаимной ориентации инерциального блока и тела системы видеоанализа. Автор работы занимался построением системы анализа механических действий человека и их классификации

для промышленного применения Подопросветов А.В. и др.. В частности такая система может применяться в области охраны труда для анализа действий человека для оценки объема движения и оценки отклонения двигательных функций от нормы (определение травмы).

На 2022 г. актуальность направления по распознаванию действий не представляет сомнений. Новые технологии становятся неотъемлемой частью как для систем безопасности рабочего процесса, так и для эффективности трудовой деятельности на промышленном предприятии. Последние достижения в данной области связаны с применением глубоких нейронных сетей с различными дополнениями, описанными в работе [44]. Также можно заметить, что количество публикаций и проработанность подходов в описанной области растет от года к году. На момент написания работы существует множество исследований с задачей определения движений человека по видео с помощью методов машинного обучения.

В работе [45] для классификации и локализации действий в видео без кадрирования авторы объединяют возможности архитектур нейронных сетей *3D – CNN* (*3D Convolutional Neural Network*) и *RNN* (*3D Recurrent Neural Network*) в единую структуру. Выходные данные из модели 3D-CNN берутся в качестве входных данных для 3D-RNN. Сеть обрабатывает 16ти кадровый клип с последовательностью признаков из видео и возвращает последовательность вероятностей классов. Модель обучается на датасете *Activity Net Challenge 2016* (640 часов видео). В ходе работы были достигнуты значения метрики *mean average precision (mAP)* $mAP=0,5874$ и $mAP=0,2237$ в задачах классификации и локализации действий соответственно.

В работе [46] рассматривается проблема точной локализации ключевых кадров действия на видео. Авторы работы описали функции потерь, которые позволили бы уменьшить число ложноположительных прогнозов. Структурированная потеря основана на наилучшем совпадении между предсказанными и размеченными началами действий. Описаны вспомогательные функции: *Matching Loss*, *Wasserstein/EMD Loss*, *Per-Frame Loss*, *Combined Loss*. Используется рекуррентная нейронная сеть для минимизации структурированных потерь с помощью градиентного спуска. Работа функций проверяется на датасетах *The Mouse Reach Dataset* (видео начинается с того, что мышь тянется к таблетке, и заканчивается, когда она ее ест) и *THUMOS'14* (большой набор видео

с размеченными действиями). Для набора данных *Mouse Reach* было обнаружено, что *Wasserstein/EMD Loss* легче оптимизируется. Однако для набора данных *THUMOS'14 Matching Loss* показала наилучшие результаты.

В работе [47] авторы исследуют проблему обнаружения действий с помощью потоковых данных о скелете. Они предлагают многозадачную рекуррентную нейронную сеть *Joint Classification-Regression* для лучшего распознавания действия и его локализации. Используя совместно задачи классификации и регрессионной оптимизации, эта сеть способна автоматически более точно находить начало и конец действий. В частности, используя преимущества подсети *deep Long Short-Term Memory (LSTM)*, предлагаемая модель автоматически фиксирует сложную долгосрочную временную динамику, которая избегает типичной конструкции скользящего окна и обеспечивает высокую вычислительную эффективность. Кроме того, подзадача регрессионной оптимизации дает возможность прогнозировать действие до его возникновения. Используется общедоступный набор данных, *G3D (Gaming Action Dataset)*, на котором, по словам авторов, модель демонстрирует многообещающую производительность.

Авторы работы [48] решают задачу распознавания последовательностей человеческих действий из видеопотока. Цель – доказательство важности обнаружения начальной точки и впоследствии предложение метода определения начала текущего действия. Метод основан на двунаправленной нейронной сети (*Bidirectional LSTM*) которая вычисляет вероятность того, что кадр станет отправной точкой, сравнивая динамику действий до и после кадра. Эксперименты на трех наборах данных (основной датасет – *Montalbano Gesture dataset*) показали, что метод может надежно обнаружить начальную точку действия, улучшая точность раннего распознавания.

OpenPos (2019) – первая система, решающая задачу определения поз для нескольких людей (англ. *Multiperson Pose Estimation*) в режиме реального времени с открытым исходным кодом [49]. Определяет 25 ключевых точек для каждого человека. Поддерживает определение не только крупных частей, но и отдельных пальцев и их движений. Для обучения использовался *CMU Panoptic Studio dataset*, состоящий из съемок людей с большого числа ракурсов в специальном куполе, оснащенный 500 камерами.

Далее рассмотрим наиболее популярные открытые обучающие выборки для классификации действий:

MSRC-Kinect12 [50], состоит из последовательности движений человека, представленных в виде положения частей человеческого тела и связанных с ними жестов, датасет содержит 594 последовательности, 719 359 фреймов (примерно 6 часов и 40 минут), собранных 30 людьми выполняющими 12 жестов, датасет был собран с помощью сенсора *kinect* и только скелетные данные доступны. *UTD-MHAD* [51], содержит 27 действий, выполненных 8 различными людьми (4 женщинами, 4 мужчинами), датасет собран с помощью сенсора *Kinect* и инерциальных датчиков. Каждый участник эксперимента выполнил каждое действие 4 раза. После удаления 3 испорченных последовательностей, датасет содержит 861 последовательность данных. Датасет содержит действия, относящиеся к следующим категориям: спортивные действия (боулинг, теннисная подача), жесты руками (нарисовать «х», нарисовать круг, нарисовать треугольник), повседневные действия (постучать в дверь, встать, сесть) и тренировочные упражнения (сгибание рук, выпады и приседания). Датасет состоит из 4 типов данных, видеоизображения, карты глубин, скелетные данные и данные, записанные с помощью *MEMS* сенсора. *UP-CV* [52], состоит из скелетных данных действий, выполняемых людьми, с целью быть настолько реалистичным, насколько это возможно. Он состоит из 10 действий, выполненных 20 людьми. Действия были выбраны таким образом, чтобы содержать обыденные действия, выполняемые в помещении и на улице (идти, хватать, почесать голову, скрестить руки, позвонить по телефону и т.д.) Длина действий варьируется между 8 и 500 кадрами, длина 90% действий находится в промежутке 23-167 кадров. *NTURGB + D120* [53] состоит из 120 категорий действий. Действия поделены на 3 основные группы: 82 относятся к ежедневным действиям (есть, двигать предметы, садится и т.д.), 12 связаны со здоровьем (приступ, обморок и т.д), 26 действий связаны со взаимодействиями людей (рукопожатие, толкание, объятия). Данные в датасете состоят из 4 разных компонентов: RGB-видео, скелетные данные, карта глубин, инфракрасное видео. Датасет был собран с помощью сенсора *Kinect*, скелетные данные в этом датасете представлены только в формате *.kinect*. *CAD – 60* [54], состоит из 60 категорий действий, содержит видео и скелетные данные в формате *.kinect*. Действия были записаны в 5 различных окружениях: офис, кухня, спальня, ванная, комната. Датасет был собран с помощью 4 участников: 2 мужчин и 2 женщин. *CAD – 120* [55], состоит из 120 категорий действий, содержит видео и скелетные данные в формате *.kinect*. Всего содержит 61 585 кадров.

Содержит разметку как конкретного действия, к примеру готовить хлопья или принимать лекарства, так и разметку общего действия: двигаться, есть, пить, открывать. *UT-Kinect* [56] содержит 10 типов действий, выполняемых людьми в домашнем окружении. Он был записан с использованием сенсора *Kinect*. Видеоизображение и карта глубин были записаны с частотой кадров 30. Разрешение карты глубин – 320x240 точек, а разрешение видеоизображения – 640x480. Действия включают в себя: идти, садиться, вставать, поднимать, нести, кидать, толкать, тянуть, волна и хлопки, выполняемые руками. Каждое действие было выполнено 10 разными людьми, датасет содержит 6220 кадров, 200 примеров всех типов действий. Длина каждого действия расположена в диапазоне от 5 до 120 кадров. *J-HMDB* [57], создан на основе *HMDB51* [58] датасета, содержащего более 5100 видеозаписей 51 различного действия, собранного из фильмов или записей с видеохостингов. Разметка этого датасета целиком по всем представленным действиям не практична, поэтому *J-HMDB* – это подгруппа видео из меньшего числа действий. Были убраны категории действий, которые содержали большое количество мимики, взаимодействий между людьми и необычных действий. В результате в датасете осталось 21 категория действий. Были убраны видео, в которых нельзя однозначно определить исполнителя. Оставшиеся видео были обрезаны таким образом, чтобы первый кадр и последний кадр соответствовали началу и окончанию действия. В результате для каждого действия осталось от 36 до 55 видео, которые содержат от 15 до 40 кадров. Всего получилось 31 838 размеченных кадров. После рассмотрения датасетов было принято решение использовать формат *.skeleton* для хранения скелетных данных, т.к. он содержит достаточное количество частей тела для сохранения качества информации и используется в большом количестве датасетов (*UTD-MHAD*, *UT – Kinect*, *NTURGB + D*, *MSRC – Kinect12*), на основе которых были разработаны модели с высокой точностью, распознающие движения людей.

Предсказание действия – это процесс классификации неполных входных данных по действию, которое еще предстоит. Одной из подзадач является предвосхищение действия (*action anticipation*), то есть его распознавание, когда еще не наблюдается ни единого момента действия, а классификация полностью основана на наблюдаемых контекстуальных подсказках. Другой – раннее предсказание действия по наблюдаемой его части (*early action prediction*). И то, и

другое является проблемами классификации, но предсказание часто требует набора данных с временными аннотациями и четкого деления между сегментом «до действия» и сегментом «во время действия» для предвосхищения действия или между «начальным действием» и «конечным действием» для раннего предсказания действия. Другое исследование [47] рассматривает проблему обнаружения действий с помощью потоковых данных о скелете. Авторы предлагают многозадачную рекуррентную нейронную сеть для лучшего распознавания действий и их локализации. Сеть способна автоматически определять начальную и конечную точки действий за счет задач классификации и регрессионной оптимизации. Использование подсети *deep Long Short-Term Memory (LSTM)* позволяет модели фиксировать сложную долгосрочную временную динамику, что повышает ее производительность. Результаты тестирования на общедоступном наборе данных *G3D (Gaming Action Dataset)* показали многообещающие результаты. В работе [48] исследуется задача распознавания последовательностей человеческих действий из видеопотока. Авторы предлагают метод обнаружения начала текущего действия на основе двунаправленной нейронной сети (*Bidirectional LSTM*), которая вычисляет вероятность начала текущего действия. Они демонстрируют важность обнаружения начальной точки и предлагают метод ее определения.

Третья задача состоит в прогнозировании траектории и ускорения каждого сегмента руки человека во время повторения целевого действия в экзоскелете для последующего решения задачи о применении дополнительной силы со стороны экзоскелета для экономии энергозатрат и помощи в совершении действия человеком. Существует область механических задач, где подобные задачи решаются при построении законов движения на основе динамических моделей для манипуляторов Ю.Ф.Голубев и др., или рассчитываются вспомогательные силы для экзоскелетов В.Е. Павловский. и др [59], [60], [61], [62], [63]. Более того часто проводят последующую оптимизацию конструкции экзоскелетов [64], [65]. Однако существенным отличием от работ, упомянутых авторов, является построение прогноза не на основе упрощенной механической модели, а на основе информации о предыдущем действии. Подход к предсказанию траектории на основе механической модели сильно осложнен выбором локальной упрощенной механической системы которая, должна быть оптимальна с точки зрения точности предсказания и времени вычисления, чтобы быть численно рассчитанной

для получения закона движения и дополнительного момента, что бы построить закон управления для приводов экзоскелета. Ранее автором был испытан подход в построении системы управления механической системой для которой была построена механическая модель с минимальным количеством допущений для которой были рассчитаны все возможные положения системы. Управление состояло в предсказании положения системы по некоторым начальными параметрам на основе известной динамической модели и их коррекции относительно неизвестных внешних параметров системы, на основе обратной связи по положению системы. Данный подход требует большой вычислительной мощности для такого сложного объекта как рука человека, поэтому был выбран метод прогнозирования основанный на *diffusion probabilistic model*. Ранее небыло упоминаний использования таких моделей для восстановления действий человека, однако такие модели широко распространены для предсказания временных рядов. Существенной сложностью для данных моделей служит подбор макропараметров прогнозирования системы. И на основе проведенных экспериментов было проверено предположение автора, что макропараметры для такой системы для одного и того же действия у разных людей не сильно изменяются, что и является существенно новым результатом в этом направлении, помимо применения модели для промышленного применения.

На момент написания работы существует множество исследований по задаче предсказания временных рядов. В данной работе рассматривается возможность предсказания динамики движений человека, на основе исторических данных о движении в условиях промышленного производства. Предполагается построить интеллектуальную систему управления экзоскелетом на основе анализа действий человека. Ранее было предложено решение задачи по предсказанию момента начала следующего действия, человеком, с точностью до 91%. В данной задаче предлагается предсказывать динамические характеристики частей тела человека во время самого движения человека используя при этом информацию о предыдущем движении благодаря модели *diffusion probabilistic model*. Она представляет собой вероятностную модель, которая генерирует новые значения временного ряда путем диффузии текущего значения через временной интервал основана на модели вариационного автоэнкодера VAE. Вместе с текущим значением ряда, в модель передается шумовой сигнал, который по мере продвижения через модель, сглаживается и интегрируется в текущее значение

временного ряда, что позволяет генерировать будущие значения с высоким качеством.

При рассмотрении аналогов описанной постановки будем рассматривать не все возможные методы для реконструкции, предсказания или определения, идентификации или классификации действий человека, а только те что сделаны на основе данных с инерциальных датчиков. Большинство исследований в этой области направлено на описание движений нижней части тела человека. Чаще других рассматривается задача предсказания походки человека, этому есть масса объяснений так как статистически значимую часть жизни средне-статистический человек проводит на ногах, включая, то что в медицине это является существенным не полностью изученным направлением. Некоторые из них [66; 67].

В области знаний по распознаванию именно действий человека довольно обширный обзор сделан в работе [41]. Однако сильный акцент сделан на использование глубокого обучения для распознавания активности на основе датчиков в таких областях как медицина, спорт, производство и т.д. Авторы считают глубокое обучение наиболее подходящим методом за счет способности обрабатывать большие объемы данных и автоматически извлекать признаки. В обзоре рассматриваются различные архитектуры глубокого обучения, такие как сверточные нейронные сети (CNN), рекуррентные нейронные сети (RNN), комбинированные архитектуры и глубокие модели с усилением обучения. Описывается, как каждая архитектура применяется для распознавания активности, их преимущества и ограничения. Также описываются различные наборы данных, используемые для обучения и оценки моделей, рассмотрены различные методы предварительной обработки данных, такие как фильтрация шума, аугментация данных и нормализация.

В ранее опубликованных статьях по этой теме обычно использовались методы машинного обучения, такие как *SVM*, *Xgboost*, *random forest*. Однако, в последние годы глубокие нейронные сети, такие как CNN, показали большую эффективность в решении задач обработки данных.

Для предсказания движения трехмерного скелета человека в ранних методах основное внимание уделялось предсказанию состояний [68; 69]. Позже были предложены модели на основе RNN, учитывающие последовательность состояний движения [67; 70—72]. используют свертки пространственных графов для

непосредственной регрессии всей последовательности. Кроме того, некоторые методы [73; 74] специально используют корреляции между суставами тела. Многомасштабные графы создаются [75], чтобы уловить различные зависимости на уровне тела. В этой работе предлагается сеть для предсказания движения, которая обеспечивает эквивариантность по евклидовому расстоянию – фундаментальное свойство, которое игнорируется предыдущими методами, чтобы получить более надежные и точные прогнозы.

Глубокие рекуррентные сети для предсказания последовательности движений людей были предложены в публикации дата сета Human 3.6M. C. Ionescu, D. Papava, V. Olaru and C. Sminchisescu. [76]

В публикации [67] представлен неплохой обзор таких методов. Работа посвящена исследованию возможностей применения рекуррентных нейронных сетей (RNN) для прогнозирования движений человека. В работе были использованы различные модели RNN, включая LSTM (Long Short-Term Memory) и GRU (Gated Recurrent Unit), а также различные методы подготовки данных и обработки сигналов, такие как нормализация и фильтрация. Для оценки качества моделей использовались различные метрики, такие как среднеквадратичная ошибка (MSE) и коэффициент корреляции Пирсона. В статье представлены важные идеи по использованию простых моделей показывающих лучшее качество чем более сложные общепризнанные.

В работе [77] предлагается сеть для предсказания движения, которая обеспечивает эквивариантность по евклидовому расстоянию - фундаментальное свойство, которое игнорируется предыдущими методами. Эквивариантность – это свойство модели или алгоритма сохранять определенные характеристики входных данных при преобразовании этих данных. Основываясь на результатах авторы продолжают совершенствовать свой подход так как он не может сравниться с так называемыми *state of art* решениями. Методы измерения затрат энергии, или мощности метаболизма, известного как респирометрия, является значимой составляющей в исследованиях, связанных с оценкой эффективности движений человека. Разнообразные методы респирометрии опираются на различные техники, такие как мониторинг сердечного ритма, калориметрия и газоанализ. В большинстве исследований, предпочтение отдается газоанализаторам, позволяющим оценивать затраты энергии на основе измерения концентраций кислорода и углекислого газа в вдыхаемом и выдыхаемом воздухе,

получаемом через специальные маски. В работе [78] представлено сравнение различных методов респирометрии, а также анализ их плюсов и минусов.

Основной **целью** диссертационной работы является построение системы, которая позволяет проводить разносторонний анализ типичных последовательностей движений человека, который необходим при построения системы управления экзоскелетом. Главной особенностью предложенной системы является широкая прикладная применимость, так как система приближена к промышленным стандартам использования. Вспомогательные цели:

1. Разработать методы определения кинематических и динамических параметров движения частей тела человека с использованием показаний микромеханических датчиков с учетом выбранной биомеханической модели тела человека. Дать теоретико-механическое обоснование предложенным методам. Реализовать разработанные методы в виде программных модулей с использованием известных программных пакетов биомеханики.
2. Разработать методы классификации и прогнозирования последовательностей движений человека с использованием информации, полученных от микромеханических датчиков и из данных видео фиксации.
3. Сравнить точность определения последовательностей движений с использованием микромеханических датчиков со способами, основанными только на видеофиксации движений в одинаковых условиях использования, по одной метрике качества.
4. Разработать систему прогнозирования движений сегментов тела человека на основе данных микромеханических датчиков, соответствующих типовым паттернам человеческих движений.

Важной особенностью предложенной системы является использование известных программных пакетов биомеханики для вычисления динамических параметров сегментов модели тела человека. Это позволило приблизить точность разработанных методов к промышленным стандартам и в дальнейшем обеспечить возможность широкого распространения разработанных методов.

Для достижения поставленных целей необходимо решить следующие **задачи**:

1. Разработать аппаратные и программные средства для сбора кинематических данных о движении сегментов тела человека, их ориентаций и

- линейных ускорений путем обработки показаний микромеханических датчиков и данных с видеокамер.
2. Создать комплексную систему анализа динамических характеристик последовательностей движений тела человека, на основе биомеханической модели тела человека.
 3. Создать методы идентификации текущих последовательностей, предсказания последующей последовательности и предсказания момента начала следующей последовательности движений человека.
 4. Создать системы предсказания ориентаций и линейных ускорений сегментов тела человека на основе данных с микромеханических датчиков.

Основные положения, выносимые на защиту:

1. Полуавтоматическая система калибровки инерциальных датчиков, которые укреплены на теле человека, в локальной системе координат и визуализации данных в модельной системе координат.
2. Способы определения динамических параметров движения сегментов тела человека и способы оценки энергозатрат и их теоретическое обоснование.
3. Алгоритм автоматического восстановления по видео данным информации трехмерных данных о движении сегментов тела человека (скелетных данных) из двух двухмерных некопланарных изображений.
4. Метод анализа последовательностей движений человека. Программно реализованы алгоритмы идентификации текущих последовательностей движений человека, предсказания последующей последовательности движений, предсказания момента начала последующей последовательности движений человека.
5. Программные средства прогнозирования кинематических характеристик сегментов тела человека во время выполнения последовательности движений.

Научная новизна: Работа посвящена анализу последовательностей движений человека в узкой, но не представленной в литературе области. В условиях работы на промышленном предприятии.

1. Предложен методы вычисления динамических параметров и оценки энергозатрат тела человека на основе показаний микромеханических датчиков и их теоретическое обоснование.

2. Предложен новый метод калибровки датчиков, расположенных на теле человека, основанный на применении статических методов и функциональных подходов калибровки.
3. Предложен новый свободный от вырождений метод восстановления пространственного положения объекта по его нескольким плоским некомпланарным изображениям.
4. Разработка и апробация классификаторов последовательностей движений из разных источников данных.
5. Прогнозирование движения сегментов тела человека на основе данных микромеханических датчиков, соответствующих типовым паттернам человеческих движений.

Научная и практическая значимость Результаты теоретических и экспериментальных исследований проведенных в данной работе, решают поставленную теоретико-практическую задачу по созданию инструмента для сбора и хранения данных, необходимого для обнаружения, определения, предсказания и прогнозирования действий людей на производстве. Результаты научного исследования применимы в развивающихся областях коллаборативной робототехники для усовершенствования взаимодействия человека с технологическим оборудованием и роботами. Проведенные научные эксперименты - позволяют оценивать производительность и контролировать безопасность рабочих на предприятии.

Степень достоверности полученных результатов обусловлена использованием классических математических методов теоретической механики, математической статистики, численных методов, математического анализа, методов оптимизации, теории вероятностей. Для аналитических и численных расчётов использованы стандартные, известные математические методы. Достоверность экспериментальных исследований подтверждена использованием современного измерительного оборудования, многократными независимыми тарировками, соответствием части результатов известным ранее.

Апробация работы. Основные результаты работы докладывались на: *Procedia Computer Science*. 14. Sep. 14th International Symposium «Intelligent Systems», INTELS 2020, XII Всероссийский съезд по фундаментальным проблемам теоретической и прикладной механики. 2019, XXXI Международная инновационная конференция молодых ученых и студентов по проблемам машинове-

дения (МИКМУС – 2019), Научная конференция: Медицина труда и промышленная экология. 2019, Ломоносов-2019. Сборник тезисов XXVI Международной научной конференции студентов, аспирантов и молодых ученых; секция «Вычислительная математика и кибернетика», Конференция: Наука и техника в дорожной отрасли. 2021., 21st International Conference "Complex Systems: Control and Modeling Problems CSCMP 2019., XXI Международной конференции., Медицина труда и промышленная экология. 2020 Extreme Robotics. 2019, Шестнадцатая Национальная конференция по искусственному интеллекту с международным участием КИИ-2018., XX Национальная конференция по искусственному интеллекту с международным участием КИИ-2022. Более того все представленные результаты были не раз проверены в проведенных лабораторных и промышленных испытаниях, а алгоритмы классификации и прогнозирования проверены методом оценки на тестовой выборке («*Out-of-Sample Evaluation*»).

Личный вклад. Все представленные в диссертации результаты получены лично автором. Далее перечислим основные подзадачи решаемые автором: Разработка алгоритмов. Разработан протокол связи устройств между собой на существующих датчиках и протоколах (*wifi* и *bluetooth*). Разработан алгоритм калибровки датчиков и браслета для сбора и визуализации данных. Разработана база данных для хранения собираемых данных, и данных для обработки методами машинного обучения. Построена биомеханическая модель тела человека. Алгоритм подсчета динамических характеристик. Построены и обучены искусственные нейронные сети для поставленных задач. Разработан алгоритм восстановления трехмерных данных о движении сегментов тела человека (скелетных данных) из двухмерных некомпланарных изображений. Построена робастная, автоматизированная, иерархическая система отбора и фильтрации скелетных данных на основе статистического анализа. Построения системы автоматизированного прогнозирования временных последовательностей кинематических характеристик будущего механического движения человека.

Разработка программного кода: Создание сценариев использования системы датчиков. Проектирование архитектуры работы и взаимодействия системы датчиков между собой. Создание пользовательского протокола работы с устройством *UI/UX* интерфейсов. Написание программного комплекса для головного устройства – браслета и датчиков. Проектирование программного комплекса

согласно протоколу по принятию данных с системы датчиков для ПК. Проектирование системы отладки системы датчиков. Проектирование системы визуализации полученных данных с функциями записи, воспроизведения, паузы, сохранения, выгрузки и тд. Проектирование системы фильтров и предобработки полученных данных (синхронизации датчиков, интерполяции дискретных данных). Написание программного комплекса для головного устройства – компьютера. Проектирование надстроек для подгрузки и выгрузке моделей машинного обучения для обработки данных. Проектирование надстроек для выгрузки данных в форматах для внешних систем в формате *xsense*, *opensim*. Проектирование генератора обучающих выборок для поставленных задач в необходимом формате: *.kinect – skeleton, coco, r3dg*. Написание программного комплекса надстроек для головного устройства. Написание программного кода для преобразование двух двумерных данных в трехмерные данные путем решение задачи оптимизации. Написание программной системы иерархической статистической фильтрации и интерполяции данных для трехмерных скелетных данных.

Публикации. Основные результаты изложены в 7 публикациях – включенных в текущий перечень ВАК.

Объем и структура работы. Диссертация состоит из введения, четырех глав, заключения и двух приложений. Полный объём диссертации составляет 169 страниц с 27 рисунками и 17 таблицами. Список литературы содержит 121 наименование.

Данная диссертация фокусируется на разработке комплексных методов анализа человеческих движений с использованием видео камер и микромеханических датчиков (MEMS). В первой главе ставится цель по разработке методов, позволяющих восстановить трёхмерные скелетные модели человека на основе обширной обработки видеоданных. Основными задачами, которые решаются в данной главе, являются точное восстановление координат различных анатомических точек тела человека на основе алгоритма OpenPose и создание алгоритмов для улучшения точности распознавания на основе геометрической информации о скелете, а также контроля качества процесса восстановления. Результаты, полученные в ходе этой главы, преобразуются в обобщенные методы и алгоритмы, которые можно применять в различных сценариях сбора данных и анализа движений.

Вторая глава посвящена анализу последовательностей движений человека на основе видеоданных. Здесь рассматриваются задачи по классификация движений с использованием данных, полученных в предыдущей главе. На основе данных классификации описываются методы прогнозирования последующих движений, включая прогноз начала следующей последовательности движений.

Для повышения точности работы программно-аппаратной системы в следующей главе представлена разработка системы сбора кинематических данных, основанная на датчиках *MEMS*. Производится описание биомеханической модели тела человека, а также описывается эксперимент по сбору данных с *MEMS* датчиков. Данная глава также включает в себя алгоритмы калибровки датчиков и практическое применение системы для анализа динамических характеристик движения, путем подсчёта мощности метаболизма мышц тела человека на основе данных с *MEMS* датчиков.

Завершающая глава фокусируется на анализе последовательностей движений человека с использованием *MEMS* датчиков. В данной главе рассматривается классификация движений человека, строится сравнительный эксперимент классификации на основе данных, полученных из разных источников. На основе более точных данных *MEMS* датчиков строится алгоритм прогнозирования ориентаций и линейных ускорений элементов модели тела человека.

Этот исследовательский труд направлен на создание комплексной методологии анализа человеческих движений.

Глава 1. Разработка системы сбора кинематических данных на основе видеоданных

1.1 Цели и задачи

Активное развитие области машинного обучения требует сбора всё большего и большего количества данных. При этом действительно впечатляющих результатов удаётся достичь только на качественных данных. Данный раздел посвящён сбору обучающих данных о последовательностях движений человека в промышленных условиях, а также разработке приложения для приведения собранных данных в наиболее универсальный формат, подходящий для применения методов машинного обучения. Используя методы машинного обучения, достигнуты значительные результаты в области распознавания и прогнозирования последовательностей движений человека. Подобные достижения имеют обширные применения в сферах реабилитационной и спортивной медицины, а также промышленной эргономики.

Цель данной главы заключается в разработке алгоритмических методов восстановления кинематики движений на основе информации, полученной с видеокамер. В рамках этой задачи также предполагается создание комплексной программно-аппаратной системы для сбора кинематических данных и проведение эксперимента в естественных условиях.

1.2 Описание эксперимента по сбору видеоданных

Далее рассматривается эксперимент, направленный на сбор данных для построения системы анализа последовательностей движений человека.

В эксперименте участвовали 6 человек, которые выполняли по 13 действий (последовательностей движений). На основе этого материала в дальнейшей задаче из разделов [2.1-2.2](#) требовалось собрать обучающую выборку для моделей

классификации, предсказания последовательности движений и предсказания начала следующей последовательности движений.

В таблице 1 представлен общий список последовательностей движений, выполняемый людьми.

В каждом видео не менее 5-7 повторений последовательностей движений, которые идут по определенному сценарию. Действия представляют собой набор движений, выполняемых в реальной жизни. На основе созданного скрипта генератора сценариев, который из 13 возможных действий выбирается n случайных действий и в случайном порядке собирает последовательность случайной длины со случайными промежутками между действиями от одной до трех секунд. Таким образом, для каждого видео получился уникальный сценарий, имеющий свою длину.

Запись выполнения работ велась с двух ракурсов, условное представление со спины и с боку от человека. Действия (последовательности движений) могли содержать поворот, при этом условный ориентир всегда менялся с одного на другой, но ни на какой другой. Человек в кадре находился в определённой трехметровой области в центрированном расположении на записи с видеокамер. Для съемки видео использовалась система из двух камер, которые фиксировались таким образом, чтобы центральные оси камер пересекались под углом, близким к прямому. Камеры были закреплены на высоте одного метра двадцати сантиметров над уровнем пола и находились в горизонтальном положении относительно земли. Центральная ось камеры (или ось оптической оси)—это вымышленная линия, которая проходит через центр объектива камеры перпендикулярно плоскости изображения. Эта линия соответствует линии, вдоль которой свет проходит через объектив и фокусируется на плоскости изображения или матрице. Когда камера фокусируется на объекте, линза объектива отклоняет падающий свет и формирует изображение на плоскости изображения, которая находится перпендикулярно центральной оси камеры.

В данной системе камер используются аналоговые видеокамеры с фоточувствительными элементами CCD (Charge-Coupled Device). CCD-сенсоры обладают высокой чувствительностью к свету и обеспечивают низкий уровень шума. Камеры имеют фиксированный фокус 22мм. Это означает, что фокусное расстояние объектива не изменяется, и они обеспечивают широкий угол обзора

Таблица 1 — Общий список действий для классификации

Название	Описание
Захват предмета	Выполняется, если человек в данный момент удерживает или прикасается к какому-либо предмету
Подъем большого предмета	Начало действия – в момент наклона, до полного разгибания прямой спины
Неправильный подъем большого предмета	Начало действия – в момент наклона, до полного разгибания спины
Подбор предмета	Начало действия с момента наклона до возвращения в исходное положение
Правильное опускание большого предмета	Действие осуществляется двумя руками, не сгибая спину (начало действия – начальный момент наклона, конец действия – выпрямившийся человек)
Неправильное опускание большого предмета	Осуществляется двумя руками, с неправильным положением спины (начало действия – начальный момент наклона, конец действия – выпрямившийся человек)
Опускание маленького предмета	Действие осуществляется руками, сгибая спину (начало действия – начальный момент наклона, конец действия – выпрямившийся человек)
Работа над столом	Выполняется из положения стоя, небольшой наклон тела выше пояса
Работа в наклоне	Выполняется из положения стоя, наклон тела выше пояса
Работа над головой	Выполняется, когда плечо горизонтально, предплечье вертикально
Работа сверху	Выполняется, когда руки вытянуты вверх
Работа на уровне груди	Выполняется, когда руки расположены в горизонте, на уровне глаз
Сидение на стуле	Человек начинает садиться из положения стоя

в 89 градусов, что позволяет охватить большую площадь наблюдения с каждой камеры.

Система связана с видеореги́стратором с помощью коаксиального кабеля. Видеореги́стратор является центральным устройством, которое принимает видеосигналы от камер и записывает их на внутренний носитель данных.

Аналоговая обработка включает в себя усиление сигнала, коррекцию цвета, управление экспозицией и контрастностью, чтобы подстроиться под существующие внешние параметры освещённости. Для передачи видеосигналов от камер к видеореги́стратору используется коаксиальный кабель. Коаксиальный кабель также обеспечивает передачу питания для камер с использованием технологии *PoweroverCoax* (PoC).

Далее на основе цифровой модуляции происходит преобразование аналоговых видеосигналов в цифровой формат. Видеореги́стратор имеет встроенные часы реального времени (*Real-Time Clock*), которые обеспечивают точное отслеживание времени и даты. Видеореги́стратор обладает средней точностью синхронизации данных из двух источников в 2 миллисекунды. Это означает, что разница между временем записи с разных камер составляет до 2 миллисекунд. Данная точность в эксперименте зависит от различных факторов. Среди основных, качество переданного сигнала и стабильность источников питания камер. Что касается характеристик мощности процессора и скорости записи, то они были взяты с запасом.

Продолжительность каждого ролика составляет 15-30 минут с частотой записи 30 кадров в секунду. В совокупности было получено около 9 часов видеопоследовательностей. Общее количество полученных кадров – 988800.

Для лучшего качества распознавания людей на видео необходимо было выполнить нескольких дополнительных условий: все конечности человека должны находиться в кадре на протяжении съёмки, а одежда человека должна быть противоположного с фоном тона.

Далее полученные видео были поделены на кадры и произведена синхронизация двух видео – техническая и экспериментальная. Техническая обеспечена архитектурой системы и преобразователем, а экспериментальная заключалась в совершении действия хлопка или моргания. Очевидно, что экспериментальная синхронизация не обеспечивает высокую точность, однако служит эвристической проверкой работоспособности технической системы синхронизации.

После этого видео отправлялись в модуль восстановления трёхмерных моделей тела человека.

После проведения сбора видеоматериалов, была произведена разметка для каждого из них. Разметка представляла собой заполнение таблицы, в которой указан номер кадра и тринадцать бинарных целевых переменных, соответствующих этим кадрам. Каждая колонка в таблице представляет отдельную целевую переменную, соответствующую определенной последовательности движений, заданной как действие. Значение 1 в столбце указывает на выполнение действия, в то время как значение 0 означает, что действие не выполняется. Заполнение целевых переменных производилось вручную путём покадрового анализа видео.

Таким образом, имеется видеоматериал с людьми, выполняющими 13 различных последовательностей движений снятых с двух различных ракурсов. Каждое действие имеет свою собственную целевую переменную в соответствии с таблицей 1. Каждая целевая переменная была названа соответствующим образом транслитом, чтобы отражать характер каждого действия, включая *zahvat_predmeta*, *podyem_bolshogo_predmeta*, *nepравilniy_pod*, *podbor_predmeta*, *opuskanie_bolshogo_predmeta*, *nepравilnoe_opuskanie*, *opuskanie_mal_predmeta*, *rabota_nad_stolom*, *rabota_v_naklone*, *rabota_nad_golovoy*, *rabota_sverhu*, *rabota_na_urovne_grudi*, и *sidenie_na_stule*.

1.3 Восстановление трёхмерной скелетной модели человека по видеоданным

1.3.1 Обработка кадров OpenPose

Видео обрабатывались системой OpenPose [79; 80], которая для каждого кадра видео возвращала координаты ключевых точек модели человека, изображенного на кадре (см. пример на рисунке 1.1). Также система может возвращать более детализированные модели рук и более детализированные ключевые точки

лица человека. В рамках рассматриваемой задачи достаточно полной модели человека, которую может определить система. Для каждой из 25 ключевых точек, расположение которых представлено на рисунке 1.2, система OpenPose возвращает 3 числа: координату X , координату Y и степень уверенности системы в правильности определения координат точки (число от 0 до 1).

Система *OpenPose* распознает ключевые точки человека по видеоданным. Ее работа делится на этапы:

1. Первым этапом идёт генерация признаков (*Featuregeneration*): На вход системе подается *RGB*-изображение, которое анализируется свёрточной нейронной сетью (*CNN*), инициализированной первыми 10 слоями *VGG-19* и донастроенной. На этом этапе генерируется набор карт признаков (F), которые затем передаются в многоуровневую *CNN*.

2. Многоуровневая *CNN* (*Multi-stageCNN*): Первый этап системы представляет собой прямую нейронную сеть, которая предсказывает набор 2D карт уверенности местоположения частей тела (по одной для каждой части тела) и набор 2D векторных полей связи частей тела (*PAFs*).

Карты уверенности кодируют уверенность модели в наличии соответствующей части тела в каждой точке изображения. Для каждой ключевой точки входного *RGB*-изображения размером $(h, w, 3)$ будет создана соответствующая карта уверенности размером (h, w) , где каждая ячейка содержит числовое значение, представляющее уверенность модели в наличии соответствующей части тела. При этом может быть несколько «пиков» значений на одной карте для каждой видимой ключевой точки в каждом человеке на изображении. Значение в каждой точке определяется гауссовым распределением вокруг реальной части тела на изображении. *PAFs* кодируют степень ассоциации между частями тела. То есть, это уверенность модели в том, что определённая пара частей тела принадлежит одному и тому же человеку и образует конечность. Для каждого пикселя в *PAF* в области, принадлежащей определённой конечности, 2D-вектор кодирует направление от одной части тела к другой, сохраняя информацию о положении и ориентации в области поддержки соответствующей конечности.

3. Анализ данных (*Parsing*): Производится набор двудольных соответствий для ассоциации кандидатов частей тела на основе предыдущих результатов. Поскольку может быть несколько человек на изображении, для каждой части тела необходимо определить, какому человеку она принадлежит.

Здесь необходимо использовать *PAFs*. *PAFs* позволяют оценить уверенность ассоциации каждой пары частей тела, то есть, принадлежат ли они одному и тому же человеку или нет. Теперь кандидаты частей тела формируют набор кандидатов на конечность. Чтобы выбрать лучших кандидатов, каждый из них оценивается с помощью вычисления линейного интеграла на соответствующем *PAF*. Этот интеграл вычисляется с использованием специальной формулы.

Таким образом, задача сводится к оптимальному соответствию между конечностями и людьми. То есть необходимо определить, какая конечность принадлежит какому человеку. Для каждой пары частей тела, например, правое плечо и правый локоть, определение оптимального соответствия сводится к задаче максимального взвешенного двудольного графа, где узлы графа – это кандидаты частей тела, а ребра – кандидаты на сегменты, с весом каждого ребра, соответствующего его оценке. Оптимальное соответствие в этой задаче можно получить с помощью венгерского алгоритма, который решает эту задачу за полиномиальное время.

Совпадение в двудольном графе – это подмножество рёбер такое, что никакие два ребра не имеют общей вершины. Максимальное взвешенное совпадение – это совпадение с максимальным суммарным весом для выбранных рёбер.

После того, как были найдены все возможные соединения конечностей, необходимо собрать их в полноценные позы для каждого человека. Для этого объединяются соединения, которые относятся к одной и той же части тела. Если соединение пытается связать две части тела, которые уже были присвоены разным людям, алгоритм распознает, что это противоречило бы связи *PAF* с более высокой уверенностью, и текущее соединение игнорируется.

В итоге система выводит 2D-координаты анатомических ключевых точек для каждого человека на изображении. Это позволяет получить информацию о расположении и позе каждого человека, опираясь на анализ видеоданных и соответствующие структуры данных, полученные в процессе выполнения алгоритма *OpenPose*.

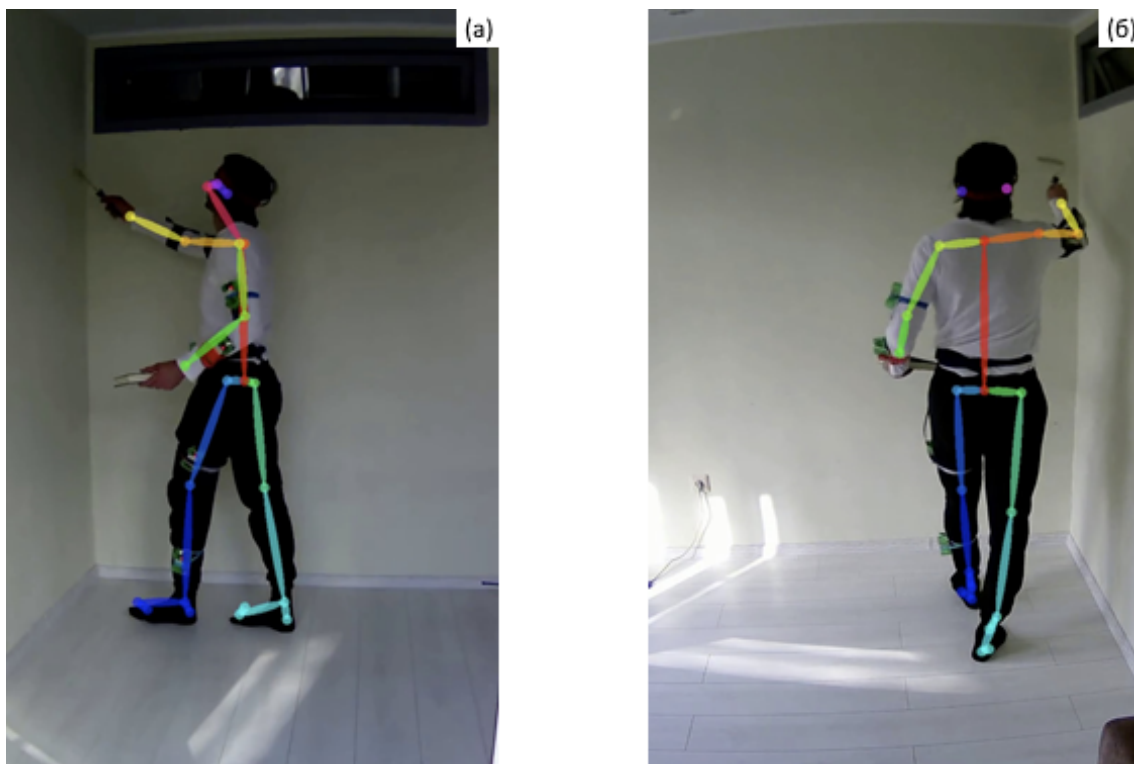


Рисунок 1.1 — Пример работы системы *OpenPose* на кадрах камер с двух ракурсов: (а) – ракурс 1, (б) – ракурс 2.

1.3.2 Постановка задачи

В ходе эксперимента записывается видео с двух синхронизированных по времени камер, расположенных под углом 90° . Затем видео разбивается на кадры. Синхронизация видео происходит автоматизированно и данные сохраняются в таблицу. Схема установки показана в двух ракурсах — вид сверху и изнутри (рис.1.3). На схеме отмечено расположение двух камер. Начало локальной системы координат в точке пересечения центральных линий объектива камер. В локальной системе координат задаем ось абсцисс на линии камеры 2, ось ординат — камеры 1.

Моделирование и визуализация алгоритма реализованы на языке Python. После сбора видеоданных они обрабатывались системой *OpenPose* — многопользовательской системой, позволяющей обнаруживать ключевые точки человеческого тела, рук, лица и ног на отдельных изображениях. В результате был подготовлен датасет, состоящий из раскодированного видео, на каждом кадре которого *OpenPose* распознаёт скелет и выводит информацию по 25 ключевым точкам в виде координат по осям X , Y и данных о степени уверенности в верном

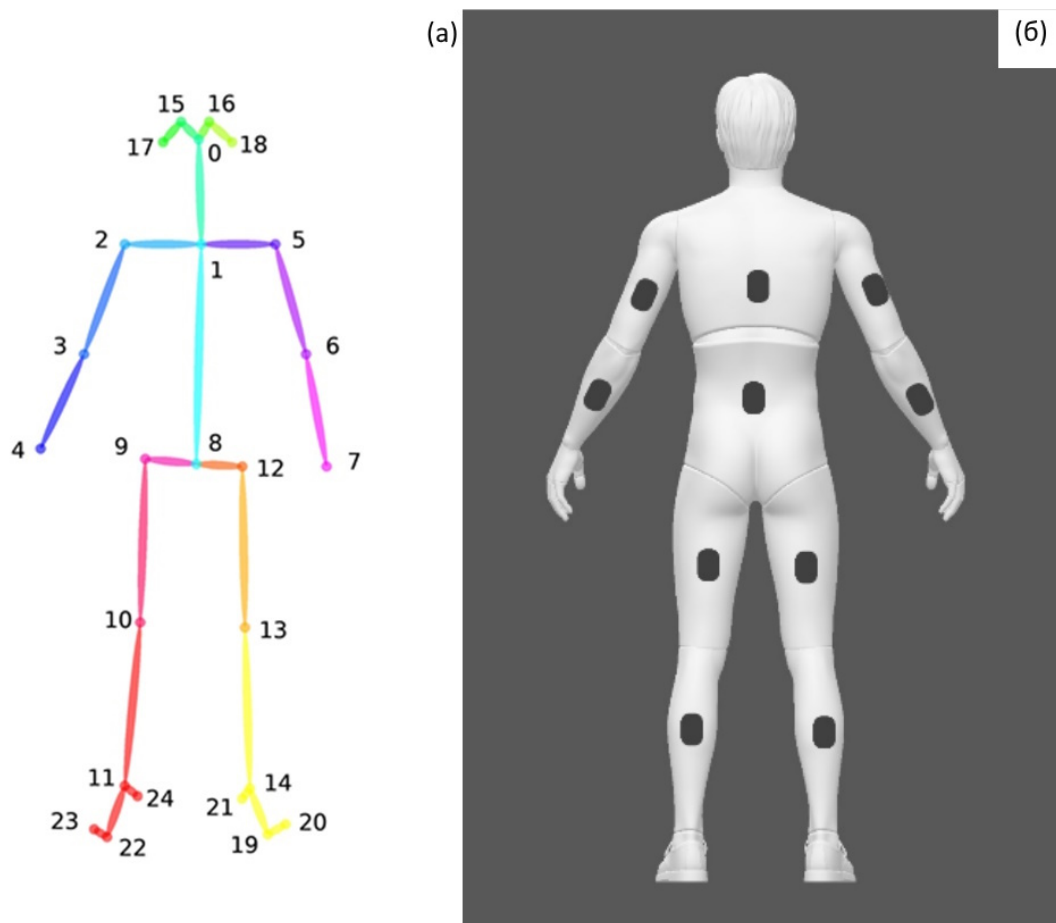


Рисунок 1.2 — (а) Номера ключевых точек, получаемых системой *OpenPose*.
 (б) Расположение датчиков IMU на теле человека.

распознавании точки. Из-за ошибок распознавания нейросетью выборка точек была отредактирована, и для алгоритма использовались только 17 значений координат точек, которые на основе статистики распознаются лучше остальных. На рис. 1.4 точки, полученные с разных изображений, выделены двумя цветами: синим цветом – точки с левой камеры (2) и красным цветом – точки с правой камеры (1).

Как видно из рисунка 1.4 задача восстановления трехмерных координат точек сводится к следующему. Если интересующая трехмерная точка проецируется в определенную позицию на наблюдаемом камерой изображении, то точка находится на некотором луче в трехмерном пространстве, поэтому точки нужно искать как пересечения этих лучей для двух камер.

Обозначим радиус-векторы искомых точек в абсолютной системе координат \vec{r}_i , $i = 1, \dots, n$, где n – количество точек, основных суставов человека. Эти векторы неизвестны, и их определение является основной задачей. Есть две камеры, описываемых радиус-вектором \vec{a}_k центра фотосенсора камеры и мат-

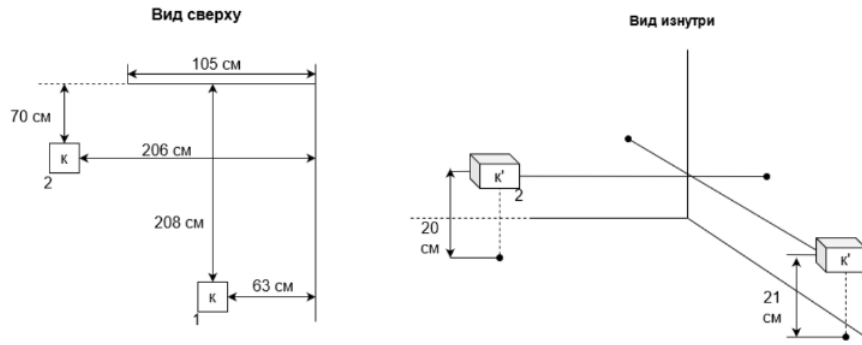


Рисунок 1.3 — Схема установки камер

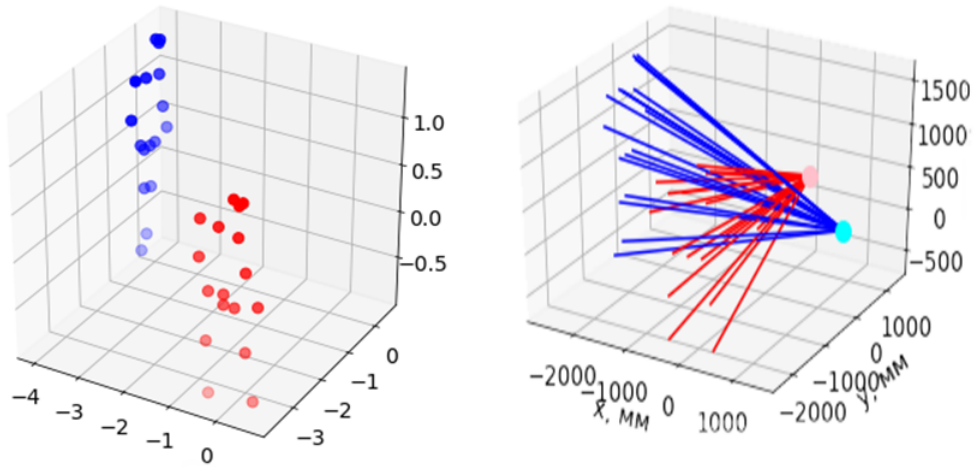


Рисунок 1.4 — Отображение точек скелета человека

рицей поворота D_k , задающей ориентацию камеры матрицей преобразования координат из связанной с камерой локальной системы координат в абсолютную, $k = 1, 2$.

Пусть координаты точки на изображении – (p_x, p_y) , координаты центра изображения в пикселях – (c_x, c_y) . Тогда в связанной с камерой локальной системе координат трехмерный луч через эту точку имеет направляющий вектор

$$\underline{u}^{loc} = \left(\frac{p_x - c_x}{w} \cdot W, \frac{p_y - c_y}{h} \cdot H, F \right),$$

где w, h – ширина и высота кадра в пикселях, W, H – ширина и высота матрицы камеры в миллиметрах, F – фокусное расстояние камеры в миллиметрах.

Получаем семейство таких лучей \vec{u}_{ki}^{loc} для обеих камер $k = 1, 2$ и всех точек скелета $i = 1, \dots, n$. С использованием матрицы поворота камеры направляющий вектор луча пересчитывается в абсолютную систему координат $\vec{u}_{ki}^{abs} = D_k \cdot \vec{u}_{ki}^{loc}$ и далее луч имеет координаты $\vec{a}_k + t \cdot \vec{u}_{ki}^{abs}$.

1.3.3 Определение координат точек

Координаты искомым точек \vec{r}_i определяются как пересечения этих лучей. Но на практике эти лучи имеют погрешности: из-за неточностей оптических систем камер, из-за неточного определения координат размещения камер \vec{a}_k и матриц поворота D_k , из-за принципиально нечеткой логики нейросети распознавания точек, поэтому фактически лучи могут быть скрещивающимися и не иметь пересечения.

В этом случае положения точек следует определять методом наименьших квадратов как точек, максимально близко расположенных к найденным лучам.

Расстояние от точки \vec{r}_i до луча $\vec{a}_k + t \cdot \vec{u}_{ki}$ можно выразить по формуле проекции вектора 1.1 используя скалярное произведение. При этом обозначим вектором \bar{u}_{ki} то как точку видит камера и как это точка существует на самом деле через искомый вектор $\bar{p}_{ik} = \vec{r}_i - \vec{a}_k$:

$$\begin{aligned}
 d_{ki} &= |\bar{p}_{ik}| \sin \alpha \\
 d_{ki}^2 &= (\bar{p}_{ik}, \bar{p}_{ik}) \sin^2 \alpha \\
 \sin^2 \alpha &= 1 - \cos^2 \alpha = 1 - \frac{(\bar{p}_{ik}, \bar{u}_{ki})^2}{(\bar{p}_{ik}, \bar{p}_{ik})(\bar{u}_{ki}, \bar{u}_{ki})} \\
 d_{ki}^2 &= (\bar{p}_{ik}, \bar{p}_{ik}) - \frac{(\bar{p}_{ik}, \bar{u}_{ki})^2}{(\bar{u}_{ki}, \bar{u}_{ki})} \tag{1.1}
 \end{aligned}$$

Обобщение алгоритма для большего числа камер

Согласно методу наименьших квадратов, координаты точек \vec{r}_i были найдены из условия минимизации суммы этих квадратов.

$$\sum_{k,i} d_{ki}^2 \rightarrow \min.$$

Заметим, что в случае двух камер эта система имеет простое явное решение – точки определяются как середины общих перпендикуляров к построенным лучам.

Данный метод был протестирован, и для повышения точности восстановления геометрии были предложены усовершенствования описанные далее.

1.3.4 Использование геометрической информации о скелете для улучшения распознавания

Пусть существует некоторое множество пар точек скелета $T = \{(i,j)\}$, для которых известны реальные расстояния между ними l_{ij} . В частности, в качестве таких точек рассматриваются концы звеньев скелета – концы геометрических представлений костей. Расстояния между ними постоянны, поскольку, как бы человек не двигался, длины его костей меняться не могут.

Тогда ограничения на расстояния между точками добавляются в полученный выше функционал МНК, его можно представить в виде

$$\alpha \sum_{k,i} d_{ki}^2 + \beta \sum_{(i,j) \in T} \left((r_i - r_j)^2 - l_{ij}^2 \right)^2 \rightarrow \min.$$

здесь α и β – варьируемые параметры-веса.

Минимизация такого функционала приведет к тому, что точки будут располагаться как можно ближе к пересечениям лучей, но одновременно на заданных расстояниях друг от друга; изменение весов α и β позволяет менять баланс между этими требованиями.

Добавление этой информации в модель эффективно решает, например, следующую проблему. В проведенных экспериментах камеры размещались под прямым углом друг к другу, но при практическом использовании этих моделей это не всегда может быть осуществимо. Иногда рабочая среда может накладывать ограничения на размещение камер. Если же камеры размещены так, что их лучи зрения почти параллельны (как глаза у человека), точность восстановления резко падает: даже небольшие отклонения по углу приведут к значительным искажениям расстояния до восстанавливаемой точки.

Естественным методом коррекции в этом случае является учет геометрических размеров: если мы знаем истинные размеры объекта, мы можем сопоставить его видимые размеры и оценить расстояние до него. И предлагаемый метод МНК с дополнительными слагаемыми в этой ситуации как раз это и сделает: восстановленные точки будут размещены как можно ближе к лучам зрения, но при этом на таком расстоянии от камер, чтобы расстояние между точками соответствовало истинному.

Недостатком этого метода является значительное повышение сложности вычислений: функционал МНК перестает быть квадратичным, его производные становятся нелинейными, и поэтому для нахождения минимума недостаточно простого решения линейной системы, требуется использовать методы численной минимизации, например, метод градиентного спуска.

Для решения этой проблемы предлагается сначала вычислять минимум функционала изначальной модели, без дополнительных, учитывающих геометрию слагаемых, этот минимум находится решением линейной системы. А далее использовать этот минимум как начальную точку для численной минимизации. Если камеры установлены достаточно точно, это начальное приближение будет близко к истинному и метод сойдется быстро, но все же позволит повысить точность.

При дальнейшем развитии модели расширенный функционал позволяет эффективно использовать информацию о точности двумерных координат точек на изображениях (уверенности распознавания), которую вычисляет OpenPose. Для этого варьируются веса в первой половине функционала:

$$\sum_{k,i} \alpha_{ki} d_{ki}^2 + \beta \sum_{(i,j) \in T} \left((r_i - r_j)^2 - l_{ij}^2 \right)^2 \rightarrow \min.$$

Здесь коэффициенты (α_{ki}) монотонно зависят от уверенности в точности распознавания ki -й точки (проекции i -й реальной точки на изображение k -й камеры), а именно α_{ki} равен этому коэффициенту уверенности.

Минимизация такого функционала приводит к тому, что точки, в которых уверенность больше, будут сильнее привязываться к своим лучам; а для точек, в которых уверенность меньше, и условие близости к лучам станет более мягким.

Таким образом, использование подобного расширенного функционала эффективно учитывает геометрическую информацию о скелете, решает ряд проблем точности распознавания в некоторых конфигурациях, и при этом учитывает информацию о достоверности распознавания точек для определения максимально правдоподобного расположения.

Наконец, отметим, что МНК можно дальше использовать и для уточнения положения камер, но данное предложение не было реализовано на практике. Прецизионно измерить их размещение и ориентацию на практике может быть затруднительно; в то же время, качество восстановления геометрии сильно за-

висит от точности, с которой известны координаты подвеса камер \vec{a}_k и матрицы поворота D_k .

Но при этом ошибки в положении камер приведут к тому, что получаемые с камер лучи перестанут пересекаться в одной точке и станут скрещивающимися. Поэтому для уточнения расположения камер можно использовать следующее требование: параметры расположения камер следует скорректировать так, чтобы эти лучи были как можно ближе к оцененным положениям точек.

На практике остаточное расстояние от лучей до точек как раз и дает значение функционала МНК в точке минимума. Поэтому для коррекции расположения камер можно дополнительно минимизировать этот функционал, используя \vec{a}_k и элементы матриц D_k в качестве переменных. Зависимость здесь также нелинейная и эту минимизацию следует проводить численно. В том случае если изначально измеренные координаты камер были достаточно точны, начальная точка будет близка к минимуму и метод сойдется быстро, то это позволит уточнить координаты.

1.3.5 Контроль восстановления скелета

Для построения трехмерной модели скелета человека введем неориентированный граф G . Он задается двумя множествами, $G = (V, E)$, где V – множество вершин (узловые точки, определенные по изображению), E – множество ребер (звенья скелета, соответствующие структуре скелета человека 1.5).

Далее были предложены дополнительные фильтрации для восстановления гладкого движения модели скелета человека. Для оценки корректности полученных трехмерных моделей скелета человека рассчитываются длины звеньев скелета и проверяется выполнение некоторых пропорций тела, а именно:

- симметричность предплечий,
- симметричность голеней,
- симметричность плеч,
- симметричность бедер.

Длины звеньев определяется по евклидову расстоянию между точками. Сравнение длин звеньев осуществляется с учетом некоторого порога ε . Накла-

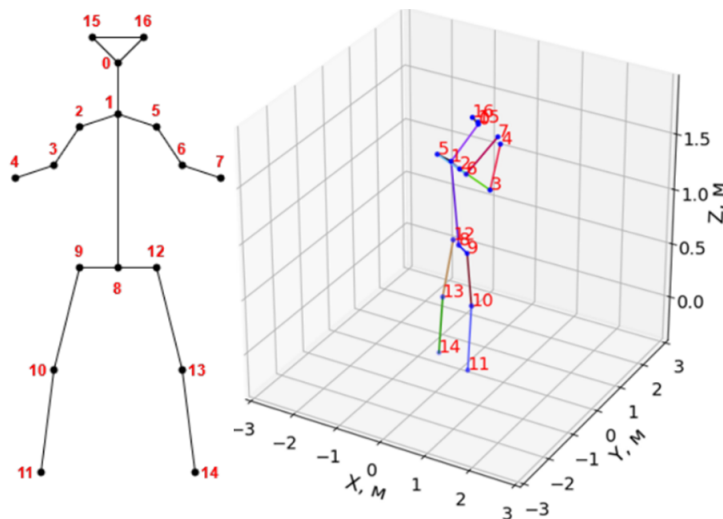


Рисунок 1.5 — 3D-модель человека, построенная по кадрам.

дывается условие, что длины двух звеньев l_1 и l_2 должны быть равны с точностью ε для данной задаче принятым в 20

$$\left| \frac{l_1}{l_2} - 1 \right| \leq \varepsilon.$$

Вычисляется отношение длин, а не их разность, так как в этом случае значение ε не связано с масштабом «скелета», который может меняться. Для каждого сравнения выбираются определенные ε , исходя из оценки достаточности данных для интерполяции. Достаточность определяется степенью корректности аппроксимации по оставшимся точкам.

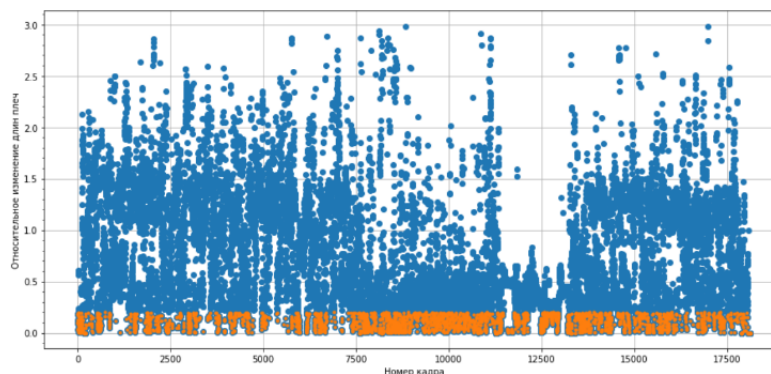


Рисунок 1.6 — Для относительного изменения длины плеч коэффициент выполнения условия оценивается в 0.2 для 5 минутного отрезка времени

На рисунке 1.6 синим цветом отображаются точки значений соотношений со всех кадров, оранжевым цветом — точки, соответствующие допустимой погрешности соотношений длин звеньев. Однако этого оказывается недостаточно.

Поэтому далее была проведена интерполяция и фильтрация значений координат точек для улучшения качества моделей 2.

Таблица 2 — Средняя оценка правдоподобности звеньев модели

Критерий	Значение	Погрешность
Предплечья	0,089	(< 0.2)
Плечи	0,002	(< 0.2)
Голени	0,128	(< 0.2)
Бедра	0,142	(< 0.2)

1.3.6 Фильтрация полученных данных

Фильтрация проводилась в два этапа. Первый исключает все плохо определенные точки – точки, для которых степень уверенности OpenPose была равна нулю. На втором этапе подбирались изменяющиеся пороговые значения для оставшихся используемых точек в совокупности с интерполяцией по оставшимся точкам для восстановления пропущенных значений. Данные операции проводились по каждому измерению.

Эвристические алгоритмы включали в себя проверку моделей скелетов на правдоподобие. Применялось три метода. Первый оценивал пропорциональное сходство распознанных частей тела; например, распознанное плечо и предплечье не должны отличаться по длине более чем на 20%, так как даже искажение перспективой не может повлиять так существенно на длину рук. Данный метод применялся как перекрестно между частями тела, так и для проверки симметрии между частями тела. Второй метод проверял совпадения определенных точек на кадре: двух плечевых суставов; двух бедренных суставов; плечевого, тазобедренного суставов, точек на грудной клетке, тазовой кости. Третий метод оценивал гладкость движений человека путем проверки скорости перемещения распознанных ключевых точек, таким образом исключая выбросы или шумящие точки выборки для последующей интерполяции.

1.3.7 Интерполяция

Для нахождения промежуточных значений координат для трехмерных моделей и обеспечения плавного движения скелетов использовалась интерполяция кубическими сплайнами (рис.1.7). Такие сплайны на каждом из частичных отрезков $[x_{i-1}, x_i]$ совпадают с кубическим многочленом и имеют на отрезке $[a, b]$ по крайней мере одну непрерывную производную $S3'(x)$.

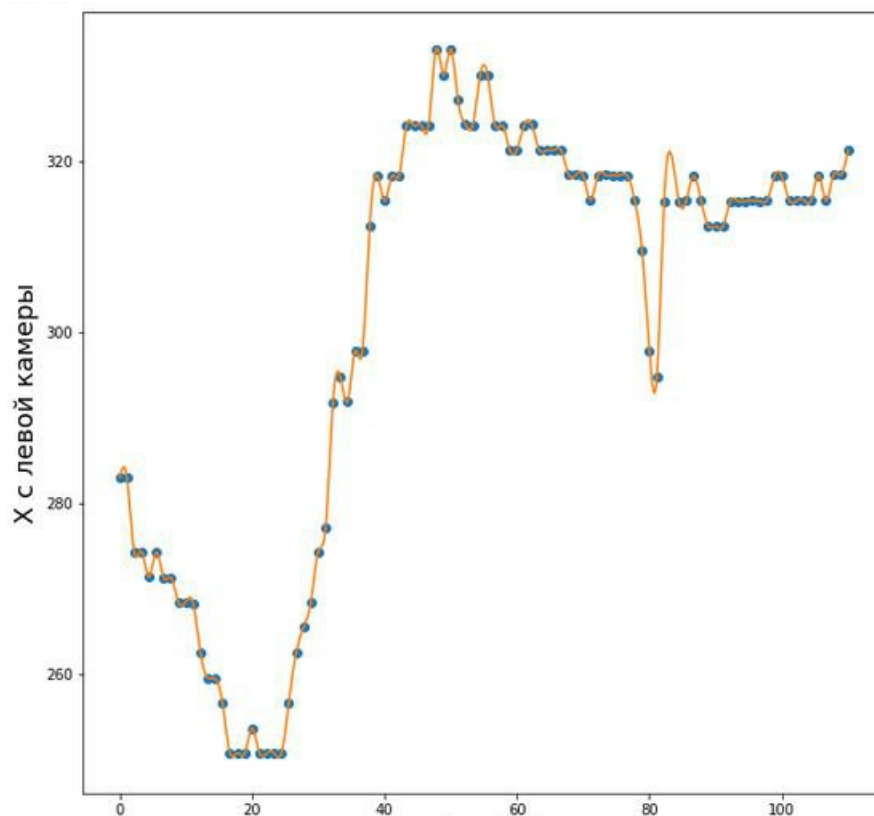


Рисунок 1.7 — График интерполированных точек после фильтрации

Выполняется построение гистограмм для всех точек и кадров по левой и правой камере 1.8.

По данной координате был выбран порог 0,2. После удаления некорректных значений точек была представлена интерполяция кубическими сплайнами отфильтрованных исходных точек.

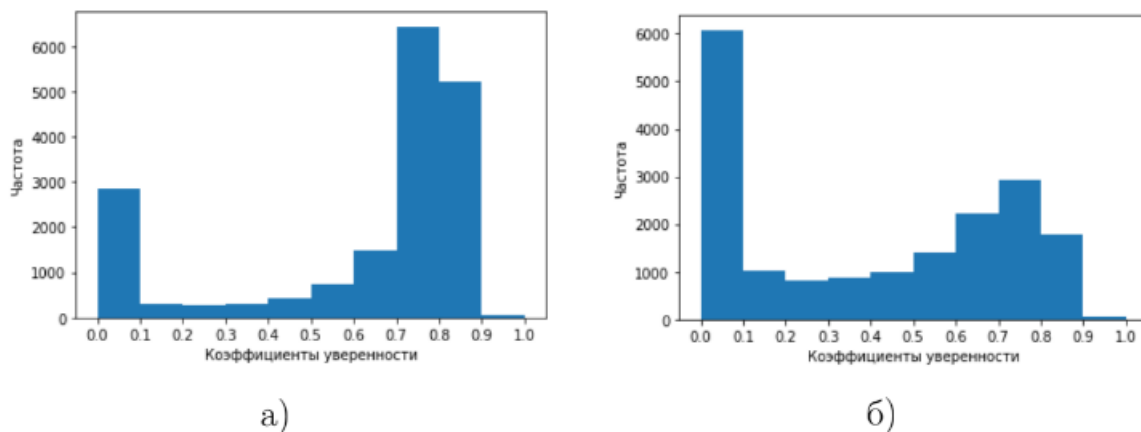


Рисунок 1.8 — Гистограмма коэффициентов уверенности точки 0.

1.4 Результаты

В результате описанных алгоритмов в первой главе, была собрана программно аппаратная система состоящая из: ПО по восстановлению трехмерных моделей по двум камерам.

На основе системы было собран модуль, который принимает видео последовательность в качестве входных данных и восстанавливает трехмерную модель скелета тела человека во времени, с учетом описанных фильтраций условий для выбора порогов и аппроксимаций.

Глава 2. Анализ последовательностей движений человека на основе видео данных

С развитием методов машинного обучения, особенно глубокого обучения, появилась возможность создавать сложные архитектуры нейронных сетей, которые обладают достаточной производительностью и способны решать широкий спектр задач в области компьютерного зрения, которые ранее не могли быть решены эффективно. В настоящее время, восстановление трехмерной модели человека на основе фото- и видеоданных, а также предсказание его действий с помощью методов машинного обучения являются очень актуальными направлениями исследований. С каждым днем возрастает число возможных сфер применения этих технологий: от взаимодействия человека с компьютерами и роботами, до анализа производительности и контроля безопасности рабочих на производстве, систем распознавания в видеонаблюдении, захвата движений в кино и компьютерных играх и многое другое.

Существует три глобальных параметра зависящие между собой: сложность построенной модели (количество обучаемых параметров), величина и качество обучающей выборки и точность получаемого результата. Отметим, что если первые два параметра приблизить к бесконечности, то при должной постановке задачи третий параметр приблизится к идеалу, однако существуют ограничения ресурсов накладываемые физическим миром, при которых описанное не достижимо. В данной главе автор ссылается на проведенный обзор существующих наборов данных, которые собраны с помощью системы kinect и являются достаточно качественными и достоверными, но не применимыми к описанным задачам. Таким образом, ставится задача собрать собственные наборы данных для решения поставленных задач. Однако, существуют ограничения в объемах и ресурсах для сбора достаточно качественных и обширных наборов данных, что приводит к использованию наиболее робастных методов на основе обзора приведенного во введении.

2.1 Классификация последовательностей движений человека

Задача состоит в разработке и обучении классификатора для автоматического определения последовательности движений человека назовем однородные последовательности движений действиями, выполняемых людьми на видео.

Введем упрощения для понимания последовательностей движений (действий) человека разделим все возможные последовательности движений на два разных типа согласно работе [81] в которой это было описано. у человека существуют Основные Действия ОД и Переходные Действия ПД, как те так и другие могут быть периодическими, в данной работе рассматриваются не периодические действия, однако это не мешает таким действиям иметь повторяющиеся элементарные шаблоны, например действие «покраска стены», «процесс прибивания ОСБ-листа», «работа с шуруповертом над головой».

Для решения поставленной задачи было выбрано строить несколько бинарных классификаторов для каждого целевого действия, используя подход «один против всех» (one vs all). «Один против всех» – это метод многоклассовой классификации, который позволяет решать задачу разделения данных на несколько классов, когда алгоритмы бинарной классификации могут быть применены только к двум классам. Суть метода заключается в последовательном обучении множества бинарных классификаторов, каждый из которых разделяет один из классов относительно остальных. В процессе обучения, для каждого классификатора, один класс рассматривается как «положительный» (класс, который мы пытаемся отличить), а остальные – как «отрицательные» классы.

После обучения всех бинарных классификаторов, для нового наблюдения каждый классификатор выдает свою оценку вероятности того, что наблюдение принадлежит соответствующему классу. Класс с наибольшей оценкой вероятности становится предсказанным классом для данного наблюдения.. Каждый классификатор обучается для различения одного целевого действия от всех остальных. Таким образом, мы получаем набор бинарных классификаторов, каждый из которых специализируется на определенном действии.

Затем производится объединение результатов этих бинарных классификаторов с помощью взвешенного голосования. Веса для каждого классификатора определяются на основе статистической встречаемости соответствующего

действия в реальной работе оператора на промышленном предприятии. Это позволяет учесть важность каждого классификатора и его специализацию на конкретном действии.

Такой подход создает комплексный ансамбль классификаторов, который способен эффективно обрабатывать различные целевые действия. Использование статистической информации о частоте встречаемости действий в реальных условиях работы оператора обеспечивает адаптивность и релевантность полученных результатов для промышленного предприятия. Это повышает производительность системы распознавания действий в реальных условиях промышленной среды.

2.1.1 Формирование датасета

Для построения модели использовались два набора данных: матрица данных с описанием и бинарные значения целевых действий (0,1). Пример разметки представлен в Приложении А 15, Эти данные были объединены в единую таблицу по номерам кадров, включающую координаты ключевых точек человека и размеченные значения целевых переменных. Для каждого видео была сформулирована таблица с 13 бинарными метками для каждого кадра, представляющими целевые переменные. По 15 лучшим видео были собраны таблицы для каждого действия, результат анализа видео последовательностей с их оценкой в Приложении 13,14,16.

Для достижения сбалансированности классов в итоговой выборке, был подсчитан каждый положительный и отрицательный пример для каждой целевой переменной. В результате получился датасет, в котором разница между числом положительных и отрицательных примеров не превышала полутора раз. Таким образом, для каждого целевого действия был сформирован свой датасет для обучения. Использование такого сбалансированного датасета повысило точность распознавания ключевых точек.

Для обеспечения хорошего качества классификатора была соблюдена сбалансированность классов, исключив лишние данные. Суммарное соотношение

положительных и отрицательных примеров в очищенных данных приведено в таблице 3.

Таблица 3 — Баланс выборок для очищенных данных

№	Целевая переменная	отрицательные/положительные примеры
1	<i>zahvat_predmeta</i>	20369/21827
2	<i>opuskanie_bolshogo_predmeta</i>	5154/3436
3	<i>rabota_v_naklone</i>	9646/6431
4	<i>rabota_nad_golovoy</i>	6940/4627
5	<i>rabota_na_urovne_grudi</i>	5865/3910
6	<i>sidenie_na_stule</i>	8067/5378

Для построения моделей бинарной классификации были отобраны шесть типов действий: 1 – захват предмета, 2 – опускание большого предмета, 3 – работа в наклоне, 4 – работа над головой, 5 – работа на уровне груди, 6 – сидение на стуле.

Однако после этого остаются кадры, на которых отсутствуют некоторые ключевые точки. В случае, если система не распознает какую-либо точку, ей присваиваются координаты и степень уверенности равные нулю. Если оставить такие координаты нулевыми, они будут иметь одинаковый масштаб с координатами точек, которые система успешно распознала. Это может негативно повлиять на качество модели, так как она может обучиться на таких точках, рассматривая их как правильные. Решение данной проблемы заключается в заполнении координат таких нераспознанных точек случайными значениями, превышающими масштаб координат «правильно распознанных точек». Заполнение координат случайными значениями большего масштаба порядка 50000 позволяет системе отличить их от точек, которые действительно присутствуют в данных и имеют смысл для обучения модели. Этот метод способствует поддержанию структуры данных и качественной обработке нераспознанных точек в процессе обучения.

2.1.2 Результаты классификации

В таблице 4 приведены результаты работы каждой модели на тренировочной и тестовой выборках.

Ассурасу (точность) и ROC-AUC (площадь под ROC-кривой) [82] – это метрики, используемые для оценки производительности классификационных моделей, особенно в задачах машинного обучения и анализа данных.

1. Ассурасу (Точность): Это простая и широко используемая метрика, которая измеряет долю правильных предсказаний модели относительно общего числа предсказаний. Она вычисляется как:

$$\text{Ассурасу} = \frac{\text{Количество правильных предсказаний}}{\text{Общее количество предсказаний}}$$

Точность может быть полезной метрикой, если классы в данных сбалансированы. Однако в случае, когда классы несбалансированы, точность может давать искаженную оценку, так как модель может успешно предсказывать больший класс и при этом иметь низкую точность для менее представленного класса.

2. ROC-AUC (Площадь под ROC-кривой) [83]: ROC (Receiver Operating Characteristic) – это график, отображающий зависимость между долей ложноположительных предсказаний и долей истинноположительных предсказаний при изменении порога классификации. ROC-AUC измеряет площадь под этой кривой и является интегральной метрикой, оценивающей общую способность модели разделять классы. Значение ROC-AUC может быть в диапазоне от 0 до 1, где значение ближе к 1 указывает на более точную модель.

Таблица 4 — Результаты двух метрик на контрольной выборке для шести целевых переменных

Название	Accuracy	AUC-ROC
Захват предмета	0.94	0.94
Работа на уровне груди	0.95	0.96
Работа в наклоне	0.95	0.95
Сидение на стуле	0.96	0.96
Работа над головой	0.96	0.96
Опускание большого предмета	0.95	0.96

2.2 Прогнозирование следующей последовательности движений

2.2.1 Набор обучающих примеров

Задача заключается в предсказании следующего действия оператора в условиях промышленного производства. В качестве входных данных используются закодированные последовательности действий человека полученные на основе эксперимента 1.4. Требуется по трем последним действиям оператора определить последующее действие. Из проведенного эксперимента были собраны цепочки действий в сумме было собрано более двухсот упоминаний каждого действия,

далее цепочки были обработаны стандартным статистическим способом. Для сравнения с общими подходами, все данные цепочек были обработаны

Требуется ограничиться знаниями о действиях полученных из предыдущего набора последовательностей движений и предсказывать одно из шести возможных действий. В рамках данной работы была собрана нейронная сеть типа LSTM (Long Short Term Memory) подробнее тип описан в работе [84].

Нейронные сети с долгой краткосрочной памятью (LSTM) представляют собой специализированный вид рекуррентных нейронных сетей (RNN), который спроектирован для эффективной работы с последовательными данными, со-

храняя долгосрочные зависимости между элементами последовательности. Они особенно полезны при обработке текстов, временных рядов и других данных, где порядок элементов играет ключевую роль.

Основная проблема стандартных RNN заключается в проблеме затухания и взрыва градиентов. LSTM были разработаны для преодоления этой проблемы, предоставляя механизмы контроля потока градиентов внутри сети. Они включают в себя три ключевых вентиля (гейта): входной, забывания и выходной.

1. Вентиль входа (Input Gate): Решает, какая информация будет добавлена в долгосрочную память. 2. Вентиль забывания (Forget Gate): Решает, какая информация будет удалена из долгосрочной памяти. 3. Вентиль выхода (Output Gate): Определяет, какую информацию из долгосрочной памяти следует использовать для выхода.

Эти вентили позволяют LSTM сохранять и использовать информацию на разных временных масштабах, что делает их особенно полезными при анализе долгосрочных зависимостей в данных.

Применение LSTM к задаче предсказания следующего действия к датасету означает, что сеть будет обрабатывать последовательность действий и их связей, учитывая окружающие действия для более точного предсказания следующего шага.

Подробно описаны параметры слоев нейронной сети на рис.2.1. В качестве входных данных используется результат работы предыдущей системы классификации, то есть номер класса действия, и ее входные данные то есть вектор из 3×25 – элементов полученных из *OpenPose*, таких вектора берется три последних действия для предсказания четвертого. Для предобработки действий используется эмбендинг слой обученный отдельно. Это позволит сети учить эффективные представления для каждого действия. Далее на входе нейронной сети многомерный вектор нормируется и подается в LSTM слой. Далее последовательность векторов пропускается через второй LSTM слой. Каждый слой LSTM будет обрабатывать входные данные и передавать свои состояния следующему слою. Количество нейронов в слое указано на изображении. В конце стоит полносвязный (Dense) слой и сформирует финальные предсказания. Последний слой размерности, равной количеству классов (в данном случае 6) с функцией активации Softmax для получения вероятностей принадлежности к

каждому классу. Выходной слой: Выходной слой даст предсказания для следующего действия в виде вероятностей. Все размерности тензоров переходов в нейронной сети выбраны с учетом получаемых кинематических данных. В данном случае n – это число обучающих примеров. Обучение нейронной сети производилось стандартным методом “обратного распространения ошибки”, с методом подбора параметров ADAM (adaptive moment estimation).

Input	seq. of n indices (range: 1..100)
1 UnitVectorLayer	matrix (size: $n \times 100$)
2 LongShortTermMemoryLayer	matrix (size: $n \times 96$)
3 LongShortTermMemoryLayer	matrix (size: $n \times 6$)
4 LinearLayer	vector (size: 6)
5 SoftmaxLayer	vector (size: 6)
6 Output	class

Рисунок 2.1 — Архитектура нейронной сети с LSTM слоями

По итогам описанных выше эксперимента, проведенного на шести испытуемых, была собрана обучающая выборка из цепочек по 3 действия с максимальным количеством примеров с каждого испытуемого на каждое из шести классифицируемых целевых действий. Итого на основе данных удалось собрать более 788×100 обучающих примеров, что подтверждается таблицы в Приложении А 17.

2.2.2 Результаты

Чтобы получить надежные, независимые от участников экспериментов результаты, используется перекрестная проверка с исключением одного испытуемого, где пропущенная часть в каждой итерации является полной коллекцией выборок от одного испытуемого. Поскольку собранные данные состоят из данных от шести испытуемых, следовательно, будет проведено шесть итерации теста, затем вычисляется среднее значение из полученных точностей для получения надежной точности.

По результатам обучения на данных на каждом из испытуемых удалось добиться 86,9%, 83,2%, 95,6%, 87,9%, 85,2%, 94,6% ожидаемых предсказаний, со средним значением 88,9%. Данные результаты говорят о том, что в действительности удастся получить обобщающее свойство модели. Объединение всех

полученных выборок привело к улучшению среднего результата предсказаний до уровня 98,7%. Однако следует отметить, что данный метод повышения точности сопровождается с кратным увеличением трудоемкости при сборе исходных данных.

2.3 Прогнозирование начала следующей последовательности движений

На основе эксперимента описанного в первой главе по сбору обучающих данных. Ставится задача о предсказании начала следующей последовательности движений человека. Для предсказания начала последовательности движений человека вручную были выбраны наиболее естественные переходы между действиями, результаты обработки видеоданных приведены в Приложении А 17.

Для выполнения основного предсказания используется структура сверточной нейронной сети (CNN). CNN (Convolutional Neural Network) парадигма которых описана в работе [85], или сверточная нейронная сеть, изначально разработана для обработки изображений, но она также может быть адаптирована для анализа временных последовательностей.

Сверточные нейронные сети имеют слои свертки, которые автоматически извлекают характерные признаки из данных. В контексте временных последовательностей, CNN может использоваться для обнаружения шаблонов или закономерностей в последовательности данных.

1. Сверточные слои (Convolutional Layers): Эти слои выполняют операцию свертки с весами, которые позволяют модели извлекать различные признаки из последовательностей. 2. Пулинг слои (Pooling Layers): Пулинг слои уменьшают размерность данных, удаляя избыточные детали и усредняя информацию о признаках. Это помогает улучшить вычислительную эффективность и обобщение модели. 3. Полносвязные слои (Fully Connected Layers): Полносвязные слои обрабатывают выходы сверточных слоев и выполняют классификацию или регрессию на основе извлеченных признаков.

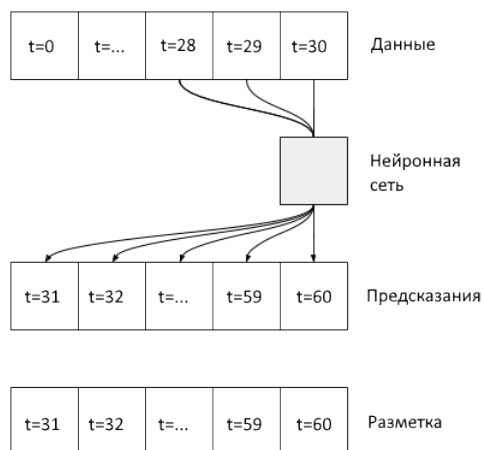


Рисунок 2.2 — Архитектура нейронной сети с CNN слоями

В данной задаче по результатам работы [86] наилучшую точность показала архитектура приведенная на рис. 2.2 содержащая 2 сверточных слоя и один полносвязный (*dense*) слой. Сверточная нейронная сеть делает прогнозы на основе истории фиксированной ширины, что может привести к лучшей производительности, чем более простые нейронные сети, поскольку она может видеть изменения, происходящие со временем.

2.3.1 Предобработка обучающих наборов данных

В ходе лабораторного эксперимента по сбору данных было отснято порядка 10 часов материала с двух ракурсов общая таблица видео файлов в Приложении А 13, 14. Все видео данные были нарезаны на 5 минутные отрезки для удобства хранения ($10 \times (60/5) = 120$ отрезков).

Датасет, использованный в исследовании, представлял собой набор извлеченных кадров из видео. На каждом из этих кадров программный инструмент OpenPose был применен для выявления характерных анатомических точек в структуре скелета. OpenPose проводил распознавание и аннотацию 25 ключевых точек на каждом кадре, предоставляя информацию о координатах каждой точки в двумерном пространстве (X, Y) и о степени уверенности в правильности данного распознавания.

Для каждой видеозаписи в исследовании была проведена оценка достоверности распознавания характерных анатомических точек скелета программой

OpenPose. Эта оценка включала в себя не только фильтрацию данных, но и применение эвристических алгоритмов, а также учет естественности действий и переходов между ними на видеозаписях.

Фильтрация данных осуществлялась двумя методами. Первый метод заключался в исключении всех точек, для которых степень уверенности, предоставленная OpenPose, равнялась нулю. Вторым методом включал в себя определение динамически изменяемых пороговых значений для сохраненных ключевых точек. Этот метод сопровождался интерполяцией значений по оставшимся точкам с учетом степени уверенности в их корректности. Данные операции выполнялись для каждой из трех осей координат.

Эвристические алгоритмы включали в себя проверку схожести модели скелета с реальной анатомией. Это включало в себя три метода. Первый метод оценивал сходство пропорций распознанных анатомических частей тела. Например, длина распознанного плеча и предплечья не должна различаться более чем на 20%, учитывая, что даже перспективные искажения не могут значительно изменить этот параметр. Вторым методом проверял соответствие определенных пар точек на кадре, таких как два плечевых сустава, два бедренных сустава, точки на грудной клетке и тазовой кости. Третьим методом оценивал плавность движений человека путем проверки скорости перемещения распознанных ключевых точек. Таким образом, исключались выбросы или зашумленные точки выборки для последующей интерполяции.

Вручную в датасете были оставлены только максимально быстрые и естественные переходы между действиями. С результатами обработки можно ознакомиться в Приложении А 17. Все подходящие данные из каждого видео объединялись в один датасет, который затем делился на слайсы – отрезки длиной в некоторое количество кадров. Затем создается окно для последовательных выборок данных, метод будет делать набор прогнозов на основе этих выборок.

Метод будет делать набор прогнозов на основе временного окна данных. Основные особенности окон ввода: ширина – число временных шагов, временной сдвиг между ними.

С целью достижения большей точности были проведена серия обучений с разными параметрами сборки обучающего набора данных 2.3. При делении данных на слайсы размером в X кадров берется по X кадров с каждой стороны от момента перехода от одного действия к другому действию и собираем обуча-

ющую последовательность из таких временных последовательностей $(\times 2) + 1$ кадров для обучения нейросети.

В качестве примера рассмотрим действие “захват предмета”, означающее, что человек на видео взаимодействует с каким-либо объектом. В ходе эксперимента было зафиксировано 71 такое действие, из которых были выделены 142 пары начала и окончания каждого действия. Учитывая, что каждое действие было записано с двух разных ракурсов, общее количество переходов составило 284. Каждый переход был взят вместе с его временной окрестностью и получил название «слайс». Размером слайса называется интервал времени, охватывающий момент перехода действия. Все подходящие данные, полученные из различных видеозаписей, были объединены в один общий датасет.

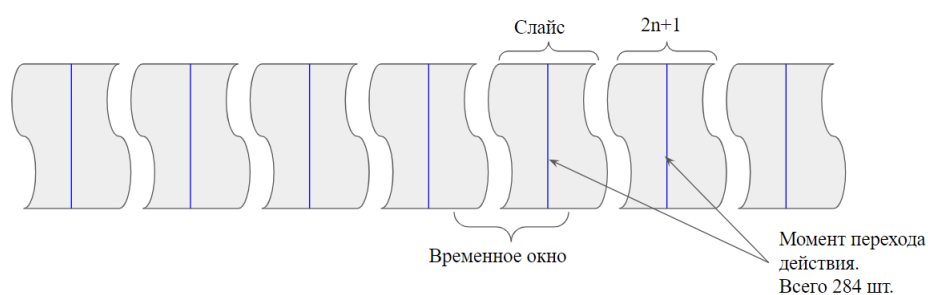


Рисунок 2.3 — Представление данных в обучающем наборе

Обучающий пример представляет собой временное окно, включающее X последовательных кадров, которые подаются на вход нейронной сети и имеют размерность (3×17) и X кадров бинарной разметки перехода целевого действия размерности $(1 \times X)$. Такое окно может содержать один или два перехода между действиями, либо не содержать их вовсе, в зависимости от длины слайса и размера окна.

Начиная с деления данных на слайсы размером в 15 кадров (половина секунды до перехода и после него), получаем датасет с 52 столбцами (17×3 – данные OpenPose, 1 – столбец разметки для предсказываемого действия “захват предмета”). Размер свертки зависит от временных интервалов.

При подаче в метод 10 кадров результаты неудовлетворительны: ошибка определения кадра начала действия 0.5715. Если рассмотреть слайсы по 30 кадров до и после каждого из переходов – переходов между действиями. При подаче по 10 кадров с целью получить на выходе предсказанные следующие 10 кадров, точность достигает 0.6279. Результаты проведенных экспериментов можно найти в таблице 5.

Таблица 5 — Результаты анализа метода

Размер слайсов	Размер окна	Точность по метрике Accuracy
15	10	0.6315
15	12	0.5932
15	15	0.6325
30	10	0.6581
30	20	0.6172
30	30	0.6301
60	20	0.8443
60	30	0.8677
60	60	0.8487

2.3.2 Результаты прогнозирования

На рисунке 2.4, представленном ниже, синими точками соединенными линией обозначены моменты времени поданные на вход для нейронной сети, зелеными точками обозначены истинные значения разметки данных, оранжевыми крестиками – предсказания модели. Перепад графика означает переход от нуля к единице или наоборот то есть совершается действие или нет. В некоторых случаях видно что нейронная сеть ожидает переход там где его нет.

При увеличении слайсов до 60 кадров (две секунды), и соответственно его размера до 121, результаты обучения улучшаются существенно. Стоит отметить, что благодаря расширению слайсов увеличивается количество примеров так как берутся все возможные примеры, где точка перехода попадает во внутренние 100 кадров из 121, или 41 для слайсов 61. Однако из-за специфики сборки данных протестировать на слайсах больше величины не удалось из за отсутствия большей части данных.

При тестировании окна по 60 кадров в обученную на слайсах длиной в 30 кадров. Средний результат предсказания шести последовательностей движений по метрике точности оказался 0.9236 рис.2.5. В первом примере модель ошиб-

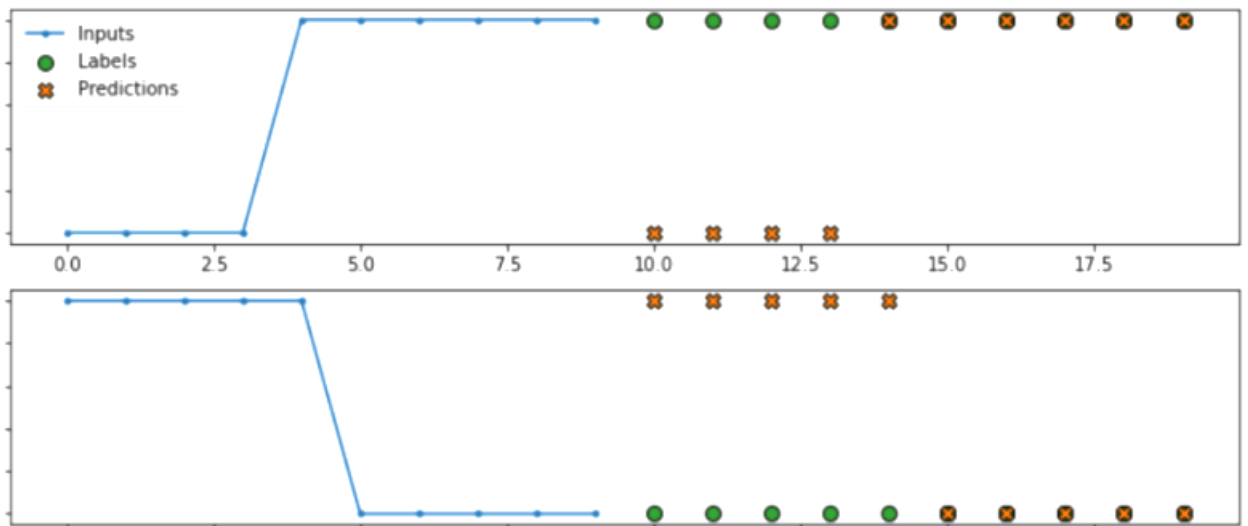


Рисунок 2.4 — Пример работы метода по определению действий по 10 кадрам лась на 16 кадров, но абсолютно верно обнаружила переход во втором примере и всего на один кадр ошиблась в третьем.

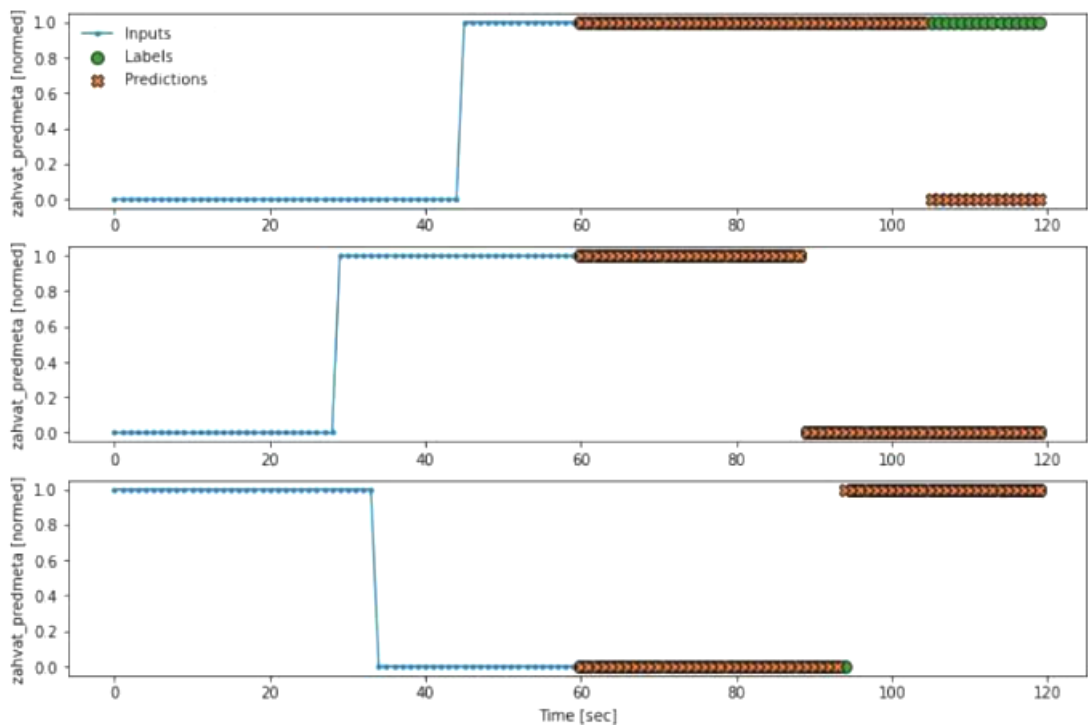


Рисунок 2.5 — Пример работы метода по определению действий по 30 кадрам

2.3.3 Заключение

По результатам тестирования нейронной сети можно сделать вывод, что самым эффективным изначальным параметром для предсказания является нарезка данных на слайсы размером в 60 кадров и подача данных в метод окном по 30 кадров.

Таким образом созданный обучающий набор данных, сверточная нейронная сеть и предложенный алгоритм в совокупности являются решением задачи по предсказанию момента начала следующего действия человека. В качестве входных данных был использован описанный эксперимент с ручной разметкой видеоданных, где моделировались возможные действия рабочего на предприятии.

В результате работы, была создана сверточная нейронная сеть, способная для шести действий предсказывать моменты окончания и начала следующей последовательности движений человека с точностью в 92%.

Глава 3. Система по сбору кинематических данных на основе датчиков MEMS

3.1 Введение

Целью главы является определение кинематических и динамических параметров движения и оценок энергетических характеристик движений человека с использованием показаний MEMS-датчиков.

(*Inertial Measurement Unit*) – это компактное устройство, основанное на *MEMS* (*Micro-Electro-Mechanical Systems*) технологии. Поэтому далее будем использовать слово сочетание MEMS датчики.

Для достижения целей были разработаны: алгоритм сбора кинематических данных с датчиков *MEMS*, алгоритм калибровки системы датчиков, алгоритм конвертации собираемых данных в известный формат kinect, а так же в известные форматы .mvpn от xsens, оценка динамических характеристик движения.

Создание инструмента включает разработку и создание набора носимых датчиков. Разработанные автором лично архитектуры и алгоритмы для работы набора датчиков включают:

1. разработанные сценарии использования системы,
2. архитектуру применения известных алгоритмов программных фильтров при сборе данных с датчиков,
3. архитектуру сбора данных с датчиков и применения известного алгоритма автокалибровки и подгрузки калибровочных файлов для датчиков,
4. архитектуру сохранения полученных данных с датчиков и верхне уровневые протоколы передачи данных основанные на известном UDP протоколе передачи данных,
5. алгоритмы работы головных устройств на которых собираются данные, хранятся и обрабатываются данные, таких устройства два: браслет и компьютер/сервер,

6. архитектуру применения известных алгоритмов обработки на головных устройствах (проверка на ошибки, фильтрация (частотные фильтры, пропуски, нули), интерполяция на сфере),
7. алгоритм восстановления трехмерной позы человека (привязка к модели человека, преобразование в скелетные данные, сохранение доп. информации о динамических характеристиках),
8. алгоритм калибровки системы датчиков на основе получаемых данных,
9. алгоритмы для визуализации получаемых данных в виде модели человека на библиотеке *vtk*,
10. алгоритм конвертации данных в известные форматы от компаний *xsens* и *microsoft*,
11. архитектуру базы данных для хранения экспериментов в базе данных *postgres*,
12. алгоритмы по оценке энергии потребляемой человеком,

Аппаратная архитектура системы сбора данных состоит из компьютера, набора из десяти IMU датчиков на технологии *MEMS*, роутера и ПО системы визуализации. Набор датчиков в своей основе использует промышленные девятиосевые сенсоры *BNO-055* от компании *Bosch*. Управление данными с датчиков выполняет двухъядерный микроконтроллер *ESP32* с низким энергопотреблением с интегрированным *Wi-Fi* и *Bluetooth* контроллерами.

В качестве собранной системы потребовалось реализовать ряд известных алгоритмов, таких как:

1. алгоритм синхронизации времен датчиков,
2. алгоритм интерполяции на сфере для кватернионов,
3. алгоритм фильтрации данных для избавления от низких и высоких частот,
4. были применены известные библиотеки обработки кватернионов.

Для осуществления сбора качественных данных и восстановления кинематики движения человека с необходимой точностью основной задачей является калибровка датчиков, а именно решение задачи отображения модели человека в соответствии с реальным миром, данный вопрос широко исследуется в литературе [87], но оптимального решения найти пока не удалось.

IMU сенсор включает в себя трёхосевой акселерометр, трёхосевой гироскоп и трёхосевой магнитометр. На основании данных, полученных с этих

устройств, вычислительный модуль датчика рассчитывает ${}_{imu}R^{glob}$ – вращение локальной системы координат (СК) F_{imu} , связанной с датчиком, относительно глобальной СК F_{glob} . Именно эти вращения считаются заданными. Так же предполагается, что система датчиков закреплена на человеке жестко, т.е. не изменяет нахождение датчика относительно части тела, но положение точки крепления на конкретной части тела неважно и неважна начальная ориентация датчика, последняя считается в ходе калибровки. С полученными данными от датчиков требуется восстановить кинематику движений человека.

Для данной задачи был разработан собственный алгоритм калибровки датчиков, подробно описанный в разделе 3.2.

3.2 Описание биомеханической модели тела человека

Для описания положения и движения сложного объекта, такого как человек, необходимым условием является построение его модели. В данном разделе предлагается ознакомиться с разработкой биомеханической модели человека по поверхностной форме тела человека.

Ранее было проведено множество исследований в области анализа движений и симуляции мускулоскелетной модели человека в области спортивных наук и медицины. Основные цели этих исследований заключаются в использовании данных о чувствительности соматических ощущений человека для улучшения движений атлетов и контроля восстановления тех, кто проходит реабилитацию подвижности. Большинство из них используют либо упрощенную модель всего тела, либо детальную модель, ограниченную малой частью тела, из-за высоких вычислительных затрат.

Вычисление прямой/обратной динамики скелетно-мышечных моделей человека с использованием данных о движении захваченных системой оптического захвата движений позволяет получать информацию для анализа кинематики и динамики. Применение этой технологии включает области спортивных наук, реабилитации и медицины. Однако системы захвата очень дороги в применении, в промышленном применении для предприятия их стоимость и полезный результат не представляют интереса. Поэтому далее рассматривается возможность

аппроксимация модели к модели возможной восстановить на основе поверхностной формы тела человека, по видео камерам, и в последствии по микроэлектромеханическим датчикам. Далее описывается создание модели тела человека.

Была разработана модель человека, показанную на рис. 1. Она состоит из скелета и мышечного комплекса. Скелет представляет собой набор костей, объединенных в определенное количество частей тела. Мышечный комплекс представляет собой набор мышц, сухожилий и связок, распространяющихся и натянутый между конкретными точками на костях тела человека. Каждый элемент этой модели описывается следующим образом.

Кость: Жесткий элемент с массой и инерцией. Мышца: Активный сжимающаяся пружина без массы. Сухожилие: Пассивная невесомая пружина, связанная с мышцей для передачи ее силы. Связка: Пассивная невесомая нить, соединяющий несколько костей для ограничения их относительного движения. Хрящ: Пассивная линейная пружина с нулевой номинальной длиной, соединяющая несколько костей и ограничивающая их относительное движение.

Предлагается рассмотреть эффективный алгоритм для вычисления прямой динамики кинематических цепей с изменяемой структурой, чтобы работать с многотелесными системами наивысшей общности. Алгоритм недавно улучшен до эффективности порядка N с использованием одного процессора и эффективности порядка $\log(N)$ с использованием множества процессоров, где N – количество тел.

Создается скелетно-мышечная модель человека так, чтобы алгоритмы вычисления динамики для многотелесных систем могли применяться. Каждый объект мышца, сухожилие и связка представляется элементом, у которой два конца и любое количество точек через которые он проходит и создает натяжение. Мышцы также могут активно менять свою длину. Предполагается, что натяжение равномерно распределено по всей длине, а точки через которые она проходит изменяют только направление элемента без трения.

Для проведения исследований механических характеристик всего тела была выполнена комбинация полигональных геометрических данных о костях, предоставленных в открытом доступе в пакете *OpenSim*, а также созданы другие полигональные данные путем восстановления геометрии из коммерческих программных продуктов модели скелета человека *AnyBody*. Обе модели основываются на стандартных антропометрических данных европейского мужчины.

С целью упрощения модели для исследования, кости были объединены между собой и было введено большее число жестких связей чем существует. Вот некоторые из них: для целей и задач мимика лица не представляет интереса, поэтому голова и челюсть были рассмотрены как одна жесткая связь. Стопа была разделена на две части: пальцы и остальную часть стопы. В результате, большинство мышц лица и коротких мышечных связок на руках и ногах были опущены. Последовательно, около 200 костей были объединены в 53 жестких элемента.

200 костей разделены на 53 группы костей, большая часть из которых соединена с соседними костями с помощью жестких связей, группы соединены между собой с помощью шаровых шарниров с тремя степенями свободы. Исключения сделаны для учета сложной подвижности, связанной с ограничениями, такими как контактные поверхности и замкнутые контуры например для плечевого сустава. Скелетная система имеет 155 степеней свободы.

Таблица 6 — Модель скелета

Наименование	Количество	Степени свободы
Кости	200	–
Группы костей	53	155
Виртуальные кости	28	168

Существует различные типы мышц, сухожилий и связок. Например, существуют крупные мышцы, такие как треугольные мышцы, длинные мышцы, обвивающие кости, а также мышцы с разветвлениями, например, бицепс. Однако их можно свести к одну из следующих вариантов:

1. простая пружина, которая соединяет начало и конец;
2. пружина, имеющей начало, вспомогательные точки и конец;
3. несколько пружин, образующими разветвление на виртуальной кости;
4. набором простых пружин;
5. комбинацией (1)–(4).

Мышечно-скелетная модель Исходя из анатомического справочника, была создана модель мышечного скелета человка. Модель охватывает большую часть мышц за исключением тех, которые находятся в области лица, шеи, кистей рук и голеностопного сустава 7.

Модель должна быть способна приблизительно воспроизводить функцию большинства типов мышц и сухожилий. Однако не учитываются следующие явления, которые могут возникнуть в реальном человеческом организме:

1. реальное добавление или удаление вспомогательных точек: движение суставов может привести к появлению новых вспомогательных точек вследствие новых контактов между мышцей и костью или удалению вспомогательной точки, когда мышца отделяется от кости;
2. контакты между элементами: не моделируются контактные силы между несколькими элементами, хотя учитываются контакты между элементами и костями с помощью вспомогательных точек.

Таблица 7 — Модель мышечного скелета

Наименование	Количество
Мышцы	366
Сухожилия	56
Связки	91
Хрящи	34

Кроме того, вводится 28 виртуальных связей, которые в сумме имеют $28 \cdot 6 = 168$ степеней свободы. Таким образом, степеней свободы кинематической цепи составляет 323. Использовалось 366, 56, 91 и 34 элемента для моделирования мышц, сухожилий, связок и хрящей соответственно. Таким образом, всего 547 элементов.

3.2.1 Упрощение модели тела человека

Очевидно что описанная модель сильно избыточная по числу мышц относительно степеней свободы и разрешение прямой или обратной задачи динамики не несет практического смысла в области знаний о работе мышц.

Следующим шагом упростим модель. Далее рассматриваются возможности для восстановления скелетно-мышечной модели человека, по поверхностной форме – по изображению человека. В контексте анатомии, поверхностная

форма описывает внешний вид человеческого тела или его частей, включая контуры, изгибы, выпуклости, ширины и длины. Изучение поверхностной формы тела помогает понять его структуру и функции.

Анатомия – это биологическая наука, которая занимается формой, положением, функцией и взаимосвязью структур в человеческом теле. Существует несколько видов анатомии, которые изучают различные аспекты структуры и функций организма:

1. Медицинская анатомия: Это основная область анатомии, которая изучает строение человеческого тела с точки зрения медицины. Она включает изучение органов, тканей, костей, мышц и других структур, а также их взаимосвязей и функций.
2. Сравнительная анатомия: Эта область анатомии сравнивает строение различных видов животных и позволяет исследовать сходства и различия между ними.
3. Развивающаяся анатомия: Этот вид анатомии изучает процессы развития организма, начиная с эмбрионального периода.
4. Функциональная анатомия: Это направление изучения анатомии, которое сосредотачивается на связи между структурой и функцией органов и систем организма. Она исследует, как различные структуры тела взаимодействуют друг с другом, чтобы выполнять определенные функции.
5. Патологическая анатомия: Эта область изучает изменения в структуре органов и тканей, вызванные различными заболеваниями и патологическими процессами.
6. Художественная анатомия: Это специализированная дисциплина, которая занимается только теми структурами, которые создают и влияют на поверхностную форму и внешний вид тела. Она помогает понять, как различные мышцы и кости взаимодействуют друг с другом при движении тела,

Три основных анатомических структуры создают поверхностную форму: Скелет, состоящий из костей и суставов, организованных в суставную структуру; Мышечная система, состоящая из сократимых мышц и неупругих сухожилий и хрящей; Подкожно-жировой слой (или жировая ткань), расположенный под кожей. Прежде чем обсуждать мышечную систему и ее влияние на поверхностную форму, вкратце упоминается о влиянии скелета. Скелет является осно-

вой всех форм поверхности. Он определяет общую форму тела и каждой из его составных частей. Скелет также влияет на поверхностную форму более непосредственно: кости создают поверхностную форму там, где кожа примыкает к костям, например, на локтях и коленях. Кости соединены в суставах, которые позволяют костям двигаться относительно друг друга. Некоторые части костей, которые в некоторых положениях не создают поверхностную форму, делают это в других положениях. Например, головки пястных костей нельзя увидеть, если рука сжата в кулак.

Среди анатомических систем, определяющих поверхностную форму, мышечная система является наиболее сложной. Мышцы располагаются рядом и слоями на верхушке костей и других мышц. Они часто пересекают несколько суставов. Мышцы обычно состоят из разных видов ткани, что позволяет некоторым частям сокращаться, а другим – нет. В зависимости от степени сокращения мышцы имеют разные формы и влияют на поверхностную форму по-разному.

Для однозначного отображения поверхностной формы тела человека будет достаточно следующей модели 8:

Таблица 8 — Скелетно-мышечная модель тела человека

Наименование	Количество	Степеней свободы
Кости	200	–
Группы костей	10	48
Виртуальные кости	6	36
Мышцы	130	
Сухожилия	45	
Связки	0	
Хрящи	34	

3.2.2 Механическая модель тела человека

Выпишем базовые уравнения движения звеньев модели скелета, для решения прямой и обратной задачи динамики. Тело человека представляется в виде древовидной скелетной системы и скелетно-мышечной системы описанной выше. Скелетно мышечная система не весома и накладывает управляющие моменты и реакции связей на возможные перемещения древовидного косного скелета состоящего из жестких тел соединенных воображаемыми суставами, которые в действительности отражены мышечной системой. Для простоты восприятия далее будет описываться закон движения скелетаной системы тела человека с суставами. Приведенная механическая модель была разработана для моделирования в системе OpenSim [18] на основе модели приведенной в работе [88]

Определения обобщенных координат и обобщенных сил для сложных ограниченных многозвенных систем. Обобщенные координаты – это минимальный набор независимых параметров, которые полностью определяют положение и ориентацию системы. В механике многозвенных систем обобщенные координаты используются для описания конфигурации системы. Например, в механической системе с несколькими звеньями, обобщенные координаты могут включать углы поворота звеньев или длины их сегментов.

Обобщенные силы – это силы, которые действуют на систему и связаны с обобщенными координатами. В механике многозвенных систем обобщенные силы позволяют описать динамику системы. Это могут быть силы тяжести, силы натяжения тросов, силы, возникающие в суставах или связях между звеньями.

Предварительные обозначения

Обозначения переменные, связанные с звеном k :

M_k – пространственная инерционная матрица $R^{6 \times 6}$,

C_k – центробежные, кориолисовы и гравитационные силы R^6 ,

ν_k – пространственная скорость R^6 ,

J_k – множество соединенных с звеном k суставов.

Переменные, связанные с суставом i :

N_{Ci} – число связей,

$N_{Fi} = 6 - N_{Ci}$ – число степеней свободы,

\dot{q}_i – скорость сустава $R^{N_{Fi}}$,

$r_{ic_i}^{\dot{}}$ – скорость дочернего звена для сустава i ,

$r_{ip_i}^{\dot{}}$ – скорость главного звена для сустава i ,

τ_i – момент для сустава $R^{N_{Fi}}$,

f_i – силы реакции связей $R^{N_{Ci}}$,

p_i – номер звена, являющегося главным элементом для сустава i ,

c_i – номер звена, являющегося дочерним элементом для сустава i ,

τ_i и f_i определены как сила и/или момент, приложенные к звену, соединенному с торцом сустава i . Другое звено получает $-\tau_i$ и $-f_i$.

Переменные, связанные с подцепочкой A : N_{LA} – количество звеньев,

N_{JA} – количество суставов,

N_{CA} – общее количество связей,

N_{FA} – общее количество степеней свободы,

ν_A – вектор, состоящий из ν_k всех звеньев $\mathbb{R}^{N_{LA}}$,

c_A – вектор, состоящий из c_k всех звеньев $\mathbb{R}^{N_{LA}}$,

M_A – блочно-диагональная матрица, состоящая из M_k всех звеньев $\mathbb{R}^{N_{LA} \times N_{LA}}$,

q_A – вектор, состоящий из q_k всех звеньев $\mathbb{R}^{N_{FA}}$,

E_A – множество внешних суставов, соединяющих подцепочку A с другими подцепочками,

L_A – множество звеньев, соединенных в цепочку A .

Далее формализуем ограничения на суставы и переменные суставов. К каждому звену прикреплена правая система координат таким образом что оси совпадают с главными осями инерций, к суставу также прикреплены две системы координат, как показано на рисунке 3.1, где звено c_i соединено с звеном p_i в сторону конечного звена. Одна система координат прикреплена к каждому звену, чья пространственная скорость обозначается как ν_k , как определено в предыдущем разделе.

Для сустава i , соединяющего два звена p_i и c_i , определяется две системы координат на рис. 3.1, каждая закреплена к одному из звеньев, пространственные скорости которых обозначаются как $r_{ip_i}^{\dot{}} \in \mathbb{R}^6$ и $r_{ic_i}^{\dot{}} \in \mathbb{R}^6$ соответственно. Направление скорости звена p_i в суставе i выбирается в противоположном направлении для согласованности с определением f_i и τ_i . Кроме того, для каждого

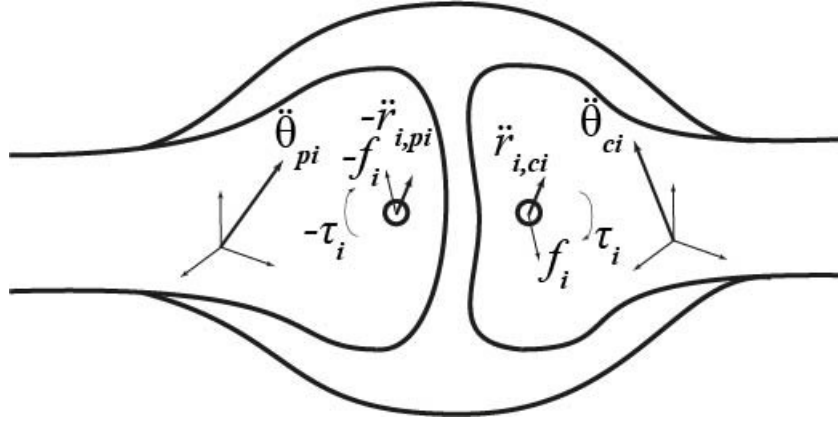


Рисунок 3.1 — Схема представления двух звеньев и сустава

сустава существует две матрицы, обозначенные как $K_{C_i} \in \mathbb{R}^{N_{C_i}}$ и $K_{J_i} \in \mathbb{R}^{N_{F_i}}$. K_{C_i} задается таким образом, что условие ограничения на суставе i уравнение связи записывается как:

$$K_{C_i} \cdot (\dot{r}_{ic_i} - \dot{r}_{ip_i}) = 0, \quad (3.1)$$

тогда как K_{J_i} задается таким образом, что скорость сустава i записывается как:

$$\dot{q}_i = K_{J_i}(\dot{r}_{ic_i} + \dot{r}_{ip_i}). \quad (3.2)$$

Предполагаем, что K_{C_i} и K_{J_i} являются постоянными для простоты, что верно для большинства практических типов суставов. Тем не менее, легко включить зависящие от времени K_{C_i} или K_{J_i} . Используя вышеуказанные величины, определяются следующие матрицы Якоби и их производные:

$$J_{ik} = \frac{\delta r_{i,k}}{\delta \phi_k},$$

$$j_{ik} = \dot{J}_{i,k} \dot{\phi}_k,$$

$$H_{ik} = K_{C_i} J_{i,k},$$

$$h_{ik} = K_{C_i} j_{ik},$$

$$H_{J_i,k} = K_{J_i} J_{i,k},$$

$$h_{J_i,k} = K_{J_i} j_{ik}.$$

Уравнение движения звена k описывается как

$$M_k \ddot{\phi}_k + C_k = \sum_{m \in J_k} (H_{m,k}^T f_m + H_{J_{m,k}}^T \tau_m), \quad (3.3)$$

$\ddot{\phi}_i$ – обобщенное ускорение звена i , $C_i(\dot{q}_i, q_i)$ – матрица Кориолисовых и центробежных сил для звена, τ_i – вектор обобщенных моментов (зависит от управляющих воздействий), где M_i – матрица инерции для звена i , J_i – матрица Якобиана для звена i , f_i – вектор сил и моментов, действующих на звено i со стороны системы. Ускорение системы координат, прикрепленной к звену k со стороны сустава i , вычисляется как

$$\ddot{r}_{ik} = J_{ik} \ddot{\phi}_k + j_{ik}. \quad (3.4)$$

Используя уравнения 3.1, 3.4, условие ограничения на суставе i , связывающее звенья p_i и c_i , описывается как

$$H_{i,c_i} \ddot{\phi}_{c_i} - H_{i,p_i} \ddot{\phi}_{p_i} + h_{i,c_i} - h_{i,p_i} = \mathbb{O}, \quad (3.5)$$

где \mathbb{O} – матрица нулей. Совместив уравнение 3.5 для всех суставов в подцепи A , получается

$$H_A \ddot{\phi}_A + h_A = \mathbb{O}, \quad (3.6)$$

где H_A – матрица размером $N_{CA} \times N_{LA}$ в следующем виде:

$$H_A = \begin{bmatrix} H_{c_1 i_1} & -H_{p_1 i_1} \\ H_{c_2 i_1} & -H_{p_2 i_1} \\ \vdots & \vdots \\ H_{c_k i_1} & -H_{p_k i_1} \end{bmatrix} \quad (3.7)$$

и h_A – вектор размером N_{CA} в следующем виде:

$$h_A = \begin{bmatrix} h_{1,c_i} - h_{1,p_i} \\ h_{2,c_i} - h_{2,p_i} \\ \vdots \\ h_{k,c_i} - h_{k,p_i} \end{bmatrix} \quad (3.8)$$

Используя H_A , глобальная форма уравнения 3.3 описывается как

$$M_A \ddot{\phi}_A + c_A = H_A^T f_A + H_{JA}^T \tau_A, \quad (3.9)$$

Уравнения 3.6, 3.7, 3.8 решаются относительно f_A как

$$f_A = S_A^{-1} \left(-H_A M_A^{-1} H_{JA}^T \tau_A + H_A M_A^{-1} c_a - h_A \right), \quad (3.10)$$

где

$$S_A = H_A M_A^{-1} (H_A)^T.$$

Вычисление уравнения 3.10 напрямую, как в большинстве коммерческих программ, приводит к алгоритму сложности $O(N^2)$, однако если использовать разреженность матрицы S_A то это позволит получить более эффективное решение сложностью $O(N)$.

3.3 Эксперимент по сбору данных с видео и MEMS

Далее будет описание другого эксперимента. В рамках лабораторного эксперимента были получены видео человека, который занимался покраской стены на протяжении продолжительного времени 8 часов, и данные с MEMS датчиков расположенных на теле человека. В данном разделе описывается два способа сбора данных. Первый способ основан на системе видеокамер описанной выше. Видео данные собирались с двух разных ракурсов (в дальнейшем обозначаются как ракурс 1 и ракурс 2) при уже описанных условиях качественных данных было собрано 5 видео длительностью в 6 минут с каждого ракурса. В совокупности было получено 1 час видеосъемок с частотой 25 кадров в секунду. Общее количество полученных кадров – 87000.

В этом же эксперименте данные записывались с датчиков *BNO055*. Датчик *BNO055* представляет собой интегрированный системный чип, который объединяет три основных датчика: трехосевой акселерометр, трехосевой гироскоп и трехосевый геомагнитный сенсор, вместе с 32-битным микроконтроллером *Cortex M0*. Все эти компоненты работают вместе для предоставления точ-

ной и совмещенной информации о движении и ориентации устройства в пространстве.

3.3.1 Описание MEMS датчиков BNO055

Датчик *BNO055* предоставляет следующие выходные данные: Кватернионы: Кватернион – это математический объект, представляющий собой четырехмерное векторное число Q выглядит следующим образом:

$$Q = q_0 + q_1i + q_2j + q_3k,$$

где q_0 – скалярная (реальная) часть кватерниона, q_1, q_2, q_3 – компоненты векторной (мнимой) части кватерниона, i, j, k – мнимые единицы, такие что $i^2 = j^2 = k^2 = ijk = -1$.

Поворот объекта в трехмерном пространстве может быть выражен с помощью кватерниона Q , который определяет угол поворота ϕ и ось вращения вектора n . Таким образом, поворот объекта представляется кватернионом:

$$Q = \cos(\phi/2) + \sin(\phi/2) \cdot (ni + nj + nk).$$

Гравитация – представляет собой вектор направления ускорения свободного падения g . Азимут – представляет собой направление, в котором направлено устройство относительно севера (N). Для обеспечения оптимальной интеграции в систему, *BNO055* оснащен цифровыми двунаправленными интерфейсами *I2C* и *UART*.

Калибровка датчика *BNO055* важна для обеспечения точности измерений. Программное обеспечение слияния данных (*sensor fusion*) датчиков выполняет калибровку акселерометра, гироскопа и магнитометра в фоновом режиме, чтобы устранить смещения. Однако для успешной автоматической калибровки необходимо выполнить предварительные шаги.

Необходимо откалибровать по отдельности акселерометр, гироскоп и магнетометр. Вот описание процедур калибровки для каждого из этих компонентов:

Калибровка акселерометра: Необходимо разместить устройство в 6 различных стабильных положениях на несколько секунд для калибровки акселерометра. Необходимо обеспечить медленное перемещение между двумя стабильными положениями. 6 стабильных положений могут быть в любом направлении, но убедитесь, что устройство лежит хотя бы один раз перпендикулярно осям x , y и z . Регистр *CALIB_STAT* позволяет проверить статус калибровки акселерометра. Калибровка гироскопа: Необходимо разместить устройство в одном стабильном положении на несколько секунд для калибровки гироскопа. Регистр *CALIB_STAT* позволяет проверить статус калибровки гироскопа. Калибровка магнитометра: Магнитометр обычно подвержен как сильным-железным (*hard-iron*), так и слабым-железным (*soft-iron*) искажениям, но в большинстве случаев присутствуют именно сильные-железные искажения. Однако при размещении датчика на плате (*PCB*) необходимо учитывать определенные предосторожности, чтобы избежать ненужных магнитных воздействий.

После калибровки датчика *BNO055* можно использовать полученный профиль калибровки для получения правильных данных об ориентации сразу после сброса питания (перед выполнением шагов калибровки, описанных выше). Однако, как только датчик переходит во внутренний режим калибровки, профиль калибровки перезаписывается новыми значениями смещения и радиуса датчика.

Для чтения и установки профиля калибровки, включая значения смещения датчиков и радиуса, необходимо перевести датчик в режим *CONFIG_MODE*. Затем можно записать соответствующие данные о смещении и радиусе и изменить режим работы датчика на режим слияния данных (*fusion mode*). Данная операция проводилась единственный раз до эксперимента в последствии использовались в качестве начальных данных записанные файлы калибровки, для каждого датчика свой.

В работе используется режим NDOF (Nine Degrees of Freedom – Девять степеней свободы), который позволяет калибровать датчик с меньшим количеством движений и обеспечивает относительно меньшее потребление энергии для калибровки магнитометра в автоматическом режиме во время работы алгоритма объединения данных с трех датчиков (*sensor fusion*). Движение для авто-калибровки может быть случайным (например, написание числа '8' в воздухе) до тех пор, пока регистр *CALIB_STAT* не покажет полную калибровку.

Для лабораторных условий и анализа движений человека, характеристики точности датчика BNO055 были определены эвристически из анализа показаний датчика в самом широком диапазоне и были зафиксированы следующие:

1. Точность измерения ускорения (акселерометр):
 - (a) Разрешение: $\pm 0.1 \text{ м/с}^2$
 - (b) Диапазон измерений: $\pm 4g$
 - (c) Чувствительность: 0.02 м/с^2
2. Точность измерения скорости вращения (гироскоп):
 - (a) Разрешение: $\pm 0.1^\circ/\text{с}$
 - (b) Диапазон измерений: $\pm 250^\circ/\text{с}$
 - (c) Чувствительность: $0.1^\circ/\text{с}$
3. Точность измерения магнитного поля (магнетометр):
 - (a) Разрешение: $\pm 0.3 \text{ }\mu\text{T}$
 - (b) Диапазон измерений: $\pm 1300 \text{ }\mu\text{T}$ (по осям x и y), $\pm 2500 \text{ }\mu\text{T}$ (по оси z)
 - (c) Чувствительность: $0.3 \text{ }\mu\text{T}$

Эти характеристики обозначают минимальные изменения, которые датчик может детектировать в каждом измеряемом параметре. В лабораторных условиях, где окружающие факторы могут быть более контролируемыми, можно ожидать, что датчик будет более точным и способным предоставлять более точные данные о движениях человека.

Однако стоит учитывать, что точность датчика также зависит от других факторов, таких как температура окружающей среды – во все датчики встроены электронные термометры, и процесс автокалибровки предполагает поправку на температуру. Однако другими факторами могут быть стабильность питания и вибрации. При анализе движений человека, особенно для более точных и сложных приложений, возможно потребуется дополнительная калибровка и компенсация ошибок, чтобы обеспечить максимально точные результаты.

3.3.2 Получение ориентаций из данных датчика BNO055

Для получения значений ориентации из сырых данных используется метод, предложенный Робертом Махони. Основная идея Махони заключается в использовании комплементарного фильтра для интегрирования данных из гироскопов и акселерометров с оценками ориентации, полученными из предыдущего состояния. Таким образом, фильтр обеспечивает стабильную и быструю оценку ориентации без необходимости сложных математических операций, таких как обращение матриц.

Кватернион ориентации обновляется путем интегрирования измерений угловой скорости гироскопа. Самый простой способ обновить кватернионы — это интегрировать дифференциальное уравнение для локальной скорости вращения.

В кинематической системе угловая скорость твердого тела в любой момент времени описывается относительно неподвижной системы отсчета, мгновенно совпадающей с системой координат тела. Таким образом, эта угловая скорость находится в системе координат твердого тела.

Используя геометрию специальной ортогональной группы, был получен пассивный комплементарный фильтр, который разделяет измерения гироскопа от реконструированного положения во входных данных. Помимо пассивного необходимо ввести явный Комплементарный Фильтр, который требует только выходы акселерометра и гироскопа, подходит для реализации на аппаратном уровне и обеспечивает хорошие оценки и вычисление смещений гироскопа в режиме реального времени.

Модели датчиков. Гироскопы измеряют угловую скорость в связанной с телом системе координат, модель ошибки которого выглядит следующим образом:

$$\Omega_y = \Omega + b + \mu \in R^3,$$

где Ω — истинная угловая скорость, b — постоянное (или медленно меняющееся во времени) смещение, и μ — добавленный шум измерения. Идеальный акселерометр измеряет мгновенное линейное ускорение \dot{v} относительно тела, за

исключением гравитационного ускорения g_0 . На практике выход a трехосевого акселерометра также имеет добавленное смещение b_a и шум μ_a .

$$a = R^T(\dot{v} - g_0) + b_a + \mu_a,$$

где $|g_0| \approx 9.8$. В квазистатических условиях этот вектор нормализуется, чтобы получить:

$$v_a = a/|a| \approx -R^T e_3,$$

где $e_3 = g_0/|g_0| = [0, 0, 1]^T$. Магнитометр предоставляет измерения магнитного поля m :

$$m = R^T(h + B_m) + \mu_b,$$

где h – магнитное поле Земли, измеренное в инерциальной системе, B_m – локальное магнитное возмущение и μ_b – шум измерения.

Интенсивность магнитного поля не имеет значения для оценки ориентации, и будет использовано только направление геомагнитного поля. Это измерение также нормализуется:

$$v_m = \frac{m}{|m|}.$$

Векторы измерений v_a и v_m используются для построения мгновенного алгебраического поворота R :

$$R \approx R_y = \arg \min_{R \in SO(3)} (\lambda_1 |e_3 - Rv_a|^2 + \lambda_2 |v_m^* - Rv_m|^2),$$

где v_m^* – справочное направление локального магнитного поля.

Соответствующие веса λ_1 и λ_2 выбираются в зависимости от относительной уверенности в выходных данных датчиков. Две степени свободы в матрице поворота определяются с использованием показаний акселерометра, и последняя степень свободы определяется с использованием магнитометра. Рассматривается система:

$$\dot{R} = R \left[\Omega \right]_{\times},$$

где $\left[\Omega \right]_{\times}$ обозначает антисимметричную матрицу $\Omega = [\Omega_X, \Omega_Y, \Omega_Z]^T$.

$$\left[\Omega \right]_{\times} = \begin{bmatrix} 0 & -\Omega_z & \Omega_y \\ \Omega_z & 0 & -\Omega_x \\ -\Omega_y & \Omega_x & 0 \end{bmatrix}.$$

Обратная операция, преобразующая антисимметричную матрицу в связанный вектор, это

$$\Omega = \text{vex}\left(\left[\Omega \right]_{\times}\right)$$

Систему можно представить в терминах кватерниона в $SO(3)$:

$$\dot{q} = \frac{1}{2}qp(\Omega),$$

где $q = (q_w, q_v) = (q_w, q_x, q_y, q_z)$ представляет собой единичный кватернион, а $p(\Omega) = (0, \Omega_X, \Omega_Y, \Omega_Z)$ представляет собой единичный чистый кватернион, ассоциированный с угловой скоростью.

Критерий ошибки. Обозначим \hat{R} как оценку матрицы вращения R в связанной с телом системе координат. Используемая ошибка оценки – это относительное вращение от связанной с телом системы координат к системе координат оценщика:

$$\tilde{R} := \hat{R}^T R.$$

Таким образом на основе анализа устойчивости Ляпунова, функции стоимости записывается:

$$E_{tr} = \frac{1}{2} \text{tr}(I_3 - \tilde{R}).$$

Цель оценки ориентации - обеспечить набор динамических уравнений для оценки $\hat{R}(t) \in SO(3)$, чтобы ошибка вращения $\tilde{R}(t) \rightarrow I_3$, что, в свою очередь, приводит к $\hat{R} \rightarrow R$.

Общая форма называется комплементарным фильтром на $SO(3)$:

$$\dot{\hat{R}} = \left[R\Omega + k_P \hat{R}\omega \right]_{\times} \hat{R},$$

где $k_P > 0$, и выражение $R\Omega + k_P \widehat{R}\Omega$ дано в инерциальной системе координат. Коррекционный коэффициент ω , полученный из ошибки \widehat{R} , можно рассматривать как нелинейную аппроксимацию ошибки между R и \widehat{R} в системе координат, связанной с \widehat{R} . Когда он не используется (то есть $\Omega = 0$), ошибка мгновенного поворота \widehat{R} будет постоянной. Если $\widehat{q} = (\widehat{q}_w, \widehat{q}_v)$ – кватернион, связанный с \widehat{R} , тогда:

$$E_{tr} = 2|\widehat{q}_\omega|^2 = 2(1 - \widehat{q}_w^2).$$

Явный комплементарный фильтр. Пусть $v_{0i} \in R^3$ обозначает набор из $n \geq 2$ известных направлений в инерциальной системе координат, где направления не коллинеарны, и $v_i \in R^3$ – соответствующие измерения. Измерения представляют собой наблюдения в системе координат, связанной с телом, фиксированных инерциальных направлений:

$$v_i = R^T v_{0i} + \mu_i,$$

где μ_i – шум процесса. Предполагается, что $|v_{0i}| = 1$ и все измерения нормализуются для $|v_i| = 1$.

Для n измерений общая стоимость становится:

$$E_{mes} = \sum_{i=1}^n k_i - \text{tr}(\widehat{R}M),$$

где $M > 0$ – положительно определенная матрица, если $n \geq 3$, или положительно полуопределенная, если $n = 2$:

$$M = R^T M_0 R,$$

$$M_0 = \sum_{i=1}^n k_i v_{0i} v_{0i}^T,$$

Веса $k_i > 0$ выбираются в зависимости от относительной уверенности в измеренных направлениях. Выразив явно комплементарный Фильтр в кватернионах, получим:

$$\dot{q} = \frac{1}{2}q \left(\Omega_y - b + k_P \cdot \omega_{mes} \right) - k_i \cdot \omega_{mes} \sum_{i=1}^n k_i v_i \cdot v_{0i}.$$

Оцененное изменение ориентации \dot{q} умножается на частоту выборки Δt для интегрирования углового перемещения, которое затем добавляется к предыдущей ориентации, чтобы получить последнюю оценку ориентации:

$$q_t = q_{t-1} + \dot{q}_t \Delta t$$

В последствии вычисленные ориентации, вектора линейных ускорений и вектора направления ускорения свободного падения сохраняются и используются для следующих целей:

1. подстановка характеристик в модель тела человека для получения динамических характеристик сегментов тела человека,
2. определение последовательности движений человека,
3. прогнозирование следующей последовательности движений,
4. прогнозирования начала следующей последовательности движений
5. прогнозирования кинематических характеристик сегментов тела человека.

3.3.3 Сбор и предобработка данных с MEMS датчиков

Подготовка данных с *MEMS* датчиков осуществлялась следующим образом. Для эксперимента использовался костюм из десяти датчиков. Каждый из датчиков состоит из микроконтроллера *ESP32*, который отвечает за опрос датчика с частотой 100 Гц и отправку данных на головное устройство, и сами датчики *BNO055* в режиме *NDOF*. Таким образом, с каждого датчика доставляется объект вращения на головное устройство. Общее количество собранных данных: 5 видео x 6 минут x 100 Гц = 150 000 кадров.

Архитектура системы для сбора и визуализации данных с датчиков, размещенных на теле человека, может быть представлена следующим образом:

1. Клиенты (Микропроцессоры *ESP32* с *Mems* датчиками *BNO055* и браслеты):
 - (а) Каждый клиент представляет собой *ESP32* микропроцессор, который собирает данные с подключенных к нему микромеханических датчиков *BNO055*.

- (b) Данные собираются периодически с заданным интервалом (ТИК), который настраивается и равен 10 мсек в данном случае.
- (c) Каждый тик собирается элементарный набор данных (БЛОК) с датчиков, состоящий из 32 байт (настраиваемый размер).
- (d) Каждый блок идентифицируется уникальным 32-битным номером, который соответствует номеру тика, в который блок был записан.

2. Буферизация:

- (a) Каждый клиент хранит данные в кольцевом буфере с размером S , который в данном случае равен 65536 блоков (настраиваемый).
- (b) Клиент запоминает полные номера буферизованных блоков, хранящихся от N до $M - 1$ (M – «голова» кольца, N – «хвост»).
- (c) При переполнении буфера ($M = N + S$) самые старые блоки отбрасываются.
- (d) Сервер хранит для каждого клиента номер X – номер последнего успешно принятого блока данных.

3. Передача данных по *WiFi* с использованием протокола *UDP*:

- (a) В каждом цикле обмена сервер отправляет широковещательный пакет-запрос всем клиентам, содержащий расписание передач и информацию о подтверждениях.
- (b) Каждый клиент, получив пакет-запрос, обрабатывает подтверждения и заменяет свое N на полученные значения X_i , если $X_i > N$.
- (c) Каждый клиент выжидает время T_i (настраиваемый параметр), после чего отправляет свои буферизованные данные в пакетах с указанием номера первого блока в пакете, номера пакета и временной метки для синхронизации (Y).
- (d) Сервер, получив данные от клиентов, сохраняет их и обновляет значения X для каждого клиента с учетом успешно принятых данных.

- (e) При потере пакетов или подтверждений сервер может повторно запросить данные, чтобы избежать потери информации.
4. Хранение данных на SD карте:
- (a) Полученные данные записываются на SD карту в порядке, в котором они приходят по WiFi.
 - (b) Каждый пакет данных содержит переменное количество блоков (N), и имеет длину $32 + 32N$ байт.

Таким образом, система позволяет собирать данные с датчиков, хранить их в кольцевом буфере для некоторого времени, передавать их по *WiFi*, сохранять на *SD* карте и подтверждать успешное получение данных серверу. Это позволит получать информацию о движении и положении человека. Далее Было реализована следующая предобработка данных:

1. Интерполяция кватернионов на сфере SQUAD (*Spherical Quaternion interpolation with QUADRATIC interpolation*). Для каждого кватерниона, который представляет ориентацию, рассчитывается его предыдущий и последующий кватернионы на основе временных меток. Затем, используя квадратичную интерполяцию между текущим, предыдущим и последующим кватернионами, получается сглаженный кватернион на сфере.
2. В *Python* для фильтрации показаний ускорений и вектора гравитации для частот 50 Гц и выше 100 Гц была использована функция фильтра нижних частот (*low-pass filter*) методом Баттерворта.
3. Для последующей визуализации данных линейных ускорений и вектора гравитации, на основе временных меток полученных с микроконтроллера аппроксимировались показания ускорений на каждом участке кусочно-полиномиальными функциями.
4. Получение всех значений на единой временной сетке, для обеспечения единой временной сетки, которая соответствует времени головного устройства (сервера), можно использовать временные метки самих данных с датчиков. Клиенты передают данные с указанием номера тика, который определяет время данных относительно старта системы в клиенте. Используя временные метки и данные о соответствии времени датчика и времени головного устройства, сервер может синхронизиро-

вать данные от разных клиентов и получить все значения на единой временной сетке для дальнейшей визуализации и анализа.

Таким образом, система предоставляет функциональность по сглаживанию, фильтрации и интерполяции данных с микромеханических датчиков, что позволяет получить более точные изменения ориентаций, линейных ускорений и вектора гравитации во времени и сохраняет их в базу данных.

3.3.4 Визуализация данных

Перед записью видео человек надевал костюм из 10 датчиков (каждому датчику соответствует одно звено в модели человека), проводилась синхронизация данных с датчиков, после чего человек совершал двухэтапный алгоритм калибровки датчиков, описанный в работе [89], который решает задачу отображения модельной системы в соответствии с реальным миром по данным *Mems* датчиков и будет описан далее. На первом этапе решается задача введения локальной системы координат связанной с моделью человека. На втором этапе решается задача определения положения датчиков относительно частей тела. После проведения алгоритма калибровки записываются видео и данные с MEMS датчиков.

Из модели получаются необходимые данные для реализации одного из трех представлений: скелетные данные в формате *.skeleton* от компании *Microsoft*, или в формате *.mvm* от компании *xsens*, или пространственной кривой движения в формате *R3DG* ([90]).

3.4 Вычисление динамических характеристик движения

3.4.1 Решения задачи динамики

Решения прямой задачи динамики

Для решения прямой задачи динамики требуется по известным обобщенным моментам, учитывая наложенные связи, посчитать необходимое движение. Существенной частью решения при этом является вычисление матрицы Якоби J . Соответственно, все решение в целом состоит из трех шагов.

Шаг первый Необходимо вычислить матрицу Якоби $J \in R^{N_l \times N_G}$

$$J = \frac{\delta l}{\delta \phi_G},$$

где N_l – число пружин и нитей в модели, N_G – число степеней свободы, $l \in R_i^N$ – это длина пружин и нитей, $\phi_G \in R^{N_G}$ – обобщенные координаты, сначала рассмотрим, J_{Li} , матрицу Якоби для i ой длины нити с учетом обобщенных координат. Матрица удовлетворяет следующему соотношению:

$$\dot{l}_i = J_{Li} \dot{\phi}_G,$$

где l_i – длина нити с номером i .

Предположим, что нить с номером i состоит из m_i точек ($m_i \geq 2$), включая промежуточные точки и начальную и конечную точки. Пусть длина между промежуточной точкой с номером j и промежуточной точкой с номером i равна $l_{i,j}$. Также пусть матрица Якоби относительно обобщенных координат будет $J_{Li,j}$. Тогда J_{Li} можно представить в виде суммы следующим образом:

$$J_{Li} = \sum_{i=0}^{m_i-2} J_{Li,j}$$

Используя $p_{i,j}$, позицию промежуточной точки с номером j для нити с номером i , $l_{i,j}$ можно описать следующим образом:

$$l_{i,j}^2 = (p_{i,j+1} - p_{i,j})^T (p_{i,j+1} - p_{i,j}).$$

Дифференцируя $l_{i,j}$ по обобщенным координатам, получим следующее уравнение:

$$J_{Li,j} = \frac{\delta l_{i,j}}{\delta \phi_G} = \frac{1}{l_{i,j}} (p_{i,j+1} - p_{i,j})^T \frac{\delta}{\delta \phi_G} (p_{i,j+1} - p_{i,j}),$$

$$J_{Li,j} = \frac{1}{l_{i,j}} (p_{i,j+1} - p_{i,j})^T (J_{pi,j+1} - J_{pi,j}),$$

где $J_{pi,j} = \frac{\delta p_{i,j}}{\delta \phi_G}$ обозначает матрицу Якоби относительно обобщенных координат, и ее можно вычислить с использованием хорошо известных методов. Сначала вычисляем J_{Li} с помощью уравнения выше, а затем получаем J , вставляя J_{Li} в i -ю строку J .

Шаг второй. На финальном шаге разрешается уравнение для обобщенных ускорений и происходит сопоставление напряжений в нитях силам и моментам в суставах. Вычисляются моменты сил в суставах $\tau_G \in R^{N_G}$ на основе заданных напряжений в нитях и пружинах f для мышц, сухожилий, связок и хрящей. Сопоставление линейно и представлено следующим уравнением:

$$\tau_G = J^T f.$$

Третий шаг

Заключается в решении уравнения 3.9, в котором известны τ_A и вычисляются $\ddot{\phi}_G$. Таким образом на основе обобщенных моментов τ_G удалось восстановить обобщенные координаты $\phi_G, \dot{\phi}_G, \ddot{\phi}_G$.

Решения обратной задачи динамики

Обратная динамика вычисляет напряжение мышц, сухожилий и связок, а также силы упругости хрящей, которые обеспечивают движение, полученное,

например, с помощью видео камер или микромеханических датчиков. Вычисление производится в два шага.

Первый шаг. Обобщенные силы (моменты сил в суставах) вычисляются по данным о движении применением стандартного обратного динамического метода Ньютона-Эйлера. Можно рассматривать механизм как кинематическую цепь с фиксированной базой, предполагая, что цепь разрезается, если в ней есть замкнутые петли, и также предполагая, что свободные суставы активно приводятся в движение. Этот подход непосредственно вытекает из основных допущений.

Второй Шаг. Расчет натяжения нитей, моделирующих мышцы, на основании найденных обобщенных моментов. Для их вычисления необходимо сначала определить обобщенные координаты, которые могут меняться во времени из-за изменения связей. Далее для расчета используется матрица Якоби J , сопоставляющая обобщенные координаты с нитями модели 8. Непосредственное сопоставление, тем не менее, провести не удастся, поскольку количество нитей существенно превосходит количество степеней свободы всей системы, и для решения этой проблемы избыточности нужно использовать методы математического программирования. В явном виде искомая связь выражается приведенным выше уравнением:

$$\tau_G = J^T f, \quad (3.11)$$

но в этом уравнении длина вектора f – сил, развиваемых нитями (мышцами), существенно превосходит длину вектора τ_G .

Напряжения волокон мышц, сухожилий и связок действуют только в направлении сокращения, в то время как напряжения пружин хрящей действуют и в обоих направлениях сокращения и растяжения. Эти характеристики представлены следующим ограничительным условием:

$$E_{mlt} f \leq 0, \quad (3.12)$$

где E_{mlt} является матрицей, извлекающей натяжения мышц, сухожилий и связок из f .

Для решения этой проблемы мы предполагаем, что у нас имеется дополнительная информация о том, как именно мышечная система человека реализует

то или иное движение. Эта информация представлена в виде функциональной зависимости

$$f_S = f_S(\phi, \dot{\phi}, s, t),$$

где f_S – это значения напряжений мышц, полученные на основе физических характеристик мышцы и ее природы. На практике эти показатели могут либо непосредственно сниматься с мышц методами электромиографии одновременно с записью данных о движении, либо браться из справочника по биомеханике мышц, где для каждой мышцы приводится информация о зависимости силы от удлинения мышцы.

После этого ставится задача нахождения f , удовлетворяющего соотношению 3.11 и при этом наиболее близкого к f_S . Формализуя это условие, требуется найти f , которое минимизирует:

$$Z = \frac{1}{2} |f - f_S(\phi, \dot{\phi}, s, t)| \quad (3.13)$$

С учетом ограничений 3.12, 3.11. Эту задачу можно решить методом квадратичного программирования (QP). А именно, решение задачи 3.13, 3.12, 3.11 можно искать как экстремум целевой функции:

$$Z = \frac{1}{2} x^T Q x + c^T x, \quad (3.14)$$

при линейных ограничениях на x (равенствах и неравенствах). Если матрица Q при этом симметрична и неотрицательно определена, задача называется выпуклой задачей квадратичного программирования и для нее известны эффективные алгоритмы. В нашем случае матрица является единичной матрицей из 3.14 и, следовательно, положительно определенной. Также можно добавить весовые коэффициенты между компонентами так, чтобы матрица оставалась положительно определенной.

Следует отметить, что построенная модель не отражает всех нюансов человеческого организма. Помимо мышц, в нем также имеются внутренние органы и ткани, осуществляющие пассивную поддержку скелета, которые нельзя описать в рамках модели нитей. В связи с этим, в некоторых случаях не удастся найти решение f , удовлетворительно приближающее найденные по записи движения обобщенные силы τ_G – удовлетворяющие основному уравнению 3.11.

3.4.2 Оценка энергии, затрачиваемой человеком

Объектом исследования данной задачи является оценка эффективности выполнения человеком некоторой заданной работы.

Целью задачи является вычисление энергоэффективности движений человека, и сравнение различных методов оценки энергозатрат человека на выполнение определенных движений.

В существующей литературе представлены разнообразные феноменологические модели для оценки энергопотребления, подробное сопоставление которых представлено в работах [91; 92] и их механические реализации [93]. Основной подход заключается в разработке механических моделей мышц, построенных на основе экспериментальных данных и биологических принципов работы мышц. Среди разнообразия моделей наиболее простой и широко используемой в биомеханическом моделировании является модель мышцы, разработанная Хиллом. Структура этой модели представлена на рисунке 3.2.

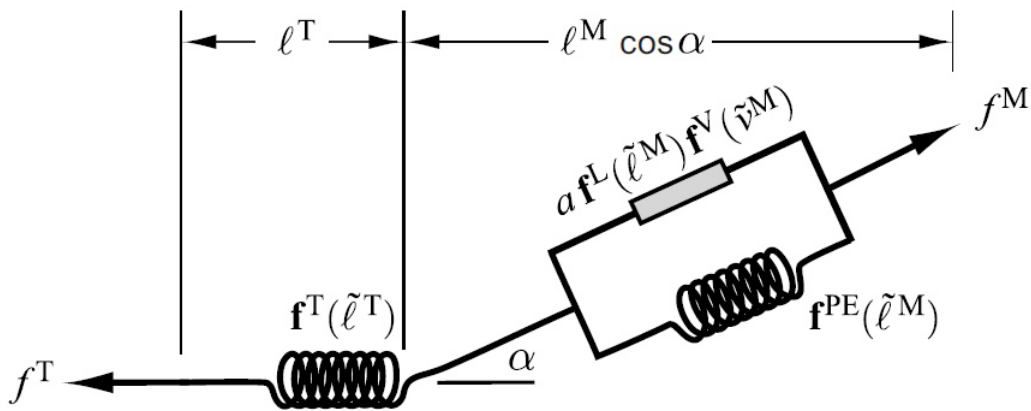


Рисунок 3.2 — Простейшая модель Хилла мышцы с сухожилием

Центральным элементом данной модели является сократительный элемент (contractile element, CE), который представляет активную силу, генерируемую мышцей. Сила, порождаемая сократительным элементом, определяется двумя законами: силы в зависимости от длины мышцы (force-length) и силы в зависимости от скорости сокращения (force-velocity). При заданных законах, выражение для силы F_M вычисляется следующим образом [94]:

$$F_M = F_{M0} \cdot a \cdot fl(\tilde{l}_M) \cdot fv(\tilde{v}_M),$$

где F_{M0} , представляет максимальную изометрическую силу (максимальную силу, которую мышца может развить при постоянной длине), а олицетворяет активацию мышцы, а fl_{l_M} и fv_{v_M} – функции изменения силы в зависимости от длины и скорости соответственно. Здесь l_M, v_M представляют нормализованные (безразмерные) величины длины мышцы и скорости сокращения соответственно. Также в модели мышцы предусмотрен пассивный элемент, который описывается нелинейно-упругой пружиной. В различных вариантах модификации данной модели могут быть реализованы более сложные комбинации последовательных и параллельных соединений пассивных элементов, таких как нелинейно-упругие пружины и демпферы.

Данная структура может варьироваться в зависимости от модификации и включать более сложные элементы. Детальное описание модели Хилла и ее различных вариаций, а также их сравнительный анализ представлены в исследованиях [91; 92].

В данной задаче рассматриваются два различных подхода: экспериментальный и основанный на моделировании. В качестве действия, для которого проводится оценка эффективности работы мышц, была выбрана покраска вертикальной стены, где основное движение рук выполняется выше уровня головы. Экспериментальный подход заключается в сравнении энергозатрат человека, определенных на основе данных пульсометра и температуры, полученных в процессе эксперимента. Вычисление энергозатрат проводится с использованием метода, описанного в работе [95].

Другое решение, предлагаемое для решения данной задачи, связано с использованием моделирования процесса выполнения работы человеком. В рамках этого подхода, в качестве биомеханической модели человека взята описанная модель для получения обобщенных сил и моментов в сочленениях, далее из предположения, что мышцы создают именно такие силы, на основе модели мышцы описанной в статье [96]. Схема данной мышцы приведена на рисунке 3.2. Для подсчета мощности метаболизма мышц при выполнении движений необходимо использовать дополнительную биомеханическую модель [97]. Она основывается на феноменологическом подходе и учитывает влияние механических параметров мышц на энергозатраты организма. Построение этой модели включает несколько ключевых шагов.

1. Действительная сила мышцы ($F_{act}(t)$): действительная сила мышцы в зависимости от времени может быть описана функцией возбуждения ($u(t)$) и максимальной изометрической силой (F_{max}):

$$F_{act}(t) = u(t) \cdot F_{max}.$$

2. Активация мышцы ($a(t)$): Активация мышцы зависит от времени и может быть описана уравнением, включающим функцию активации и характеристики сокращения мышц:

$$a(t) = A \cdot (1 - \exp(-b \cdot F_{act}(t))),$$

где A и b – параметры мышцы, $F_{act}(t)$ – активная сила мышцы в момент времени t .

3. Интеграл для энергии метаболизма (W_{met}): Мощность метаболизма связана с интегралом от квадрата активной силы по времени:

$$W_{met} = k_{met} \cdot \int_0^T a(t)^2 \cdot F_{max}^2 dt,$$

где k_{met} – коэффициент, связанный с производительностью метаболизма.

4. Максимальная сила (F_{max}): Максимальная изометрическая сила мышцы зависит от её размеров и характеристик:

$$F_{max} = \alpha \cdot PCSA \cdot F_o,$$

где α – коэффициент, связанный с анатомическими параметрами; $PCSA$ – площадь поперечного сечения мышцы, F_o – нормированная максимальная изометрическая сила.

Эти формулы описывают основные принципы работы модели [97], позволяющей оценивать мощность метаболизма мышц на основе механических параметров и активации мышц при выполнении различных движений.

Эксперимент

Для эксперимента был выбран мужчина в возрасте 26 лет (обозначенный как А) с весом 77 кг (W_b) и ростом 190 см (H_b). Используемые обозначения для констант соответствуют предыдущему исследованию [95]. Площадь поверхности тела (BSA) была вычислена согласно формуле Дюбуа и Дюбуа:

$$BSA = 0,007184 \cdot W_b^{0,425} \cdot H_b^{0,725} = 2,0428 \text{ м}^2.$$

Эксперимент был проведен при стандартной комнатной температуре. Во время эксперимента была измерена пульсовая частота с частотой 0,2 Гц при помощи спортивного браслета Mi Band 4. Также была замерена начальная и конечная температура тела. В эксперименте использовались три инерциальных датчика IMU BNO055, записывавших кинематические характеристики движения. Один датчик располагался на животе на уровне пояса, второй – на плече, третий – на предплечье. По завершении эксперимента, инерциальные данные были обработаны специальной программой и преобразованы в файл движения (.mot), который использовался для последующего моделирования в программе OpenSim.

Процесс эксперимента следующий: после минутной калибровки инерциальных датчиков, в течение 5 минут испытуемый выполнял движения правой рукой, аналогичные движениям при покраске стены. При этом разрешалось двигаться корпусу, но не нижней части тела. После этого периода, испытуемый отдыхал, находясь в сидячем положении в течение 4 минут, после чего вновь измерялся пульс.

Оценка энергозатрат по стандарту ISO8996

В данной исследовательской работе используется методология оценки мощности метаболизма, представленной в статье [95]. В указанной статье представлена формула для вычисления мощности метаболизма (MP), измеряемой в Ваттах:

$$M_{HR} = M_0 + \frac{MWC - M_0}{HR_{max} - HR_0} \cdot (HR - HR_0), \quad (3.15)$$

где M_0 и HR_0 – значения мощности метаболизма и частоты сердцебиения в состоянии покоя, MWC – максимальная работоспособность в Ваттах, HR_{max} – максимальная частота сердцебиения. Частота сердечного ритма измеряется в ударах в минуту (уд/мин).

Далее представлены формулы и разъяснения, касающиеся расчета каждого компонента уравнения 3.15. Максимальная работоспособность (MWC) или базовая скорость выделения тепла представляет собой величину максимальной потребляемой организмом мощности кислорода, выраженной в Ваттах. Формула для вычисления MWC также заимствована из статьи [95]:

$$MWC = (19.45 - 0.133 \cdot Y) \cdot LBM = 1001.6 \text{ Вт.}$$

Символ Y обозначает возраст испытуемого, а LBM – мышечную массу тела (lean body mass), рассчитанную согласно формуле из той же статьи:

$$LBM = (1.08 - \frac{W_b}{80} \cdot H_b^2) \cdot W_b = 62.63 \text{ кг.}$$

Здесь H_b представляет собой рост испытуемого в метрах. Максимальная частота сердцебиения HR_{max} определяется согласно уравнению:

$$HR_{max} = 208 - 0.7 \cdot A = 189.8 \text{ уд/мин.}$$

Частота сердечного ритма в состоянии покоя $HR_0 = 75$ уд/мин. Значение мощности метаболизма в состоянии покоя M_0 предполагается равным

$$M_0 = 60 \text{ Вт} \cdot m^2 \cdot BSA = 122.6 \text{ Вт.}$$

Согласно статьям [95; 98], в уравнение 3.15 следует внести скорректированное значение пульса HR , исключив при этом тепловую компоненту. Для оценки тепловой ошибки пульс после четырех минут отдыха вычитался из пульса H_0 , замеренного перед началом работы в состоянии покоя.

Затем строилась линейная зависимость тепловой ошибки, начиная с нулевого значения при начале работы и достигающая ранее вычисленного значе-

ния тепловой ошибки к концу работы. После этого измеренные значения HR в процессе работы корректировались на тепловую ошибку в соответствующий момент времени и подставлялись в уравнение 3.15. Полученные результаты представлены на графиках на рисунке 3.3.

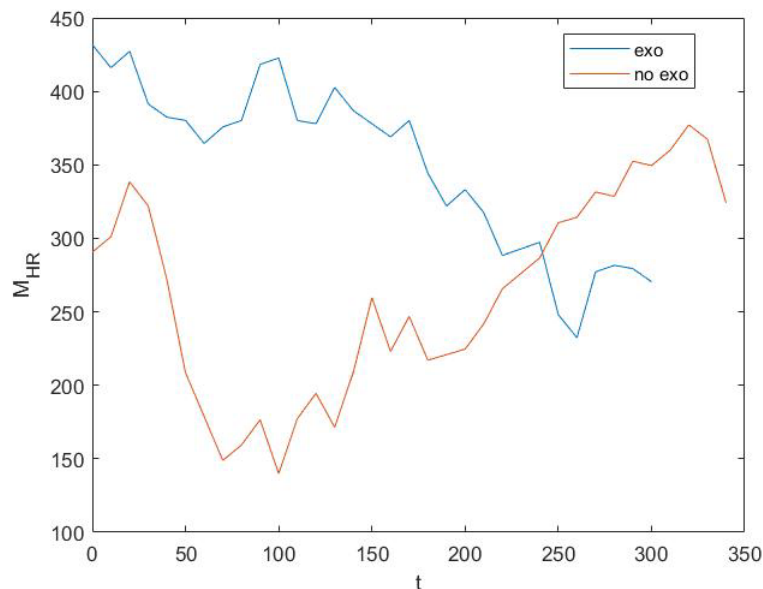


Рисунок 3.3 — Результаты расчета мощности метаболизма по стандарту *ISO8996* для пятиминутного эксперимента

Такое поведение мощности метаболизма можно объяснить особым эффектом, характерным для исследований энергозатрат при использовании экзоскелета. Более подробно про эффективность использования можно прочесть в работе [99]. Этот эффект заключается в следующем: на начальных этапах использования экзоскелета мощность метаболизма выше, чем при отсутствии экзоскелета. Это объясняется непривычностью работы с экзоскелетом и дополнительной нагрузкой на организм. Однако с течением времени, при привыкании к экзоскелету, мощность метаболизма при использовании экзоскелета становится ниже, чем при аналогичном движении без него, так как происходит быстрое утомление в последнем случае.

Оценка энергозатрат на основе модели

Второй метод решения поставленной задачи заключается в проведении моделирования. В задаче использовалось всего три *MEMS* датчика, это позволяет восстановить движения для трех частей тела.

Для анализа была взята описанная модель, однако большинство мышц и костей были зафиксированы, что бы исключить сильное влияние накапливаемой ошибки при численных вычислениях. Модель содержит 33 мышцы верхней части тела и 3 основных твердых части тела грудина с позвоночником, плечевая кость, лучевая и локтевая кости. Следует отметить, что модель является односторонней, то есть лишь правая сторона грудины учитывается при моделировании всех костей и мышц.

Процесс моделирования начинается с преобразования данных, полученных в ходе эксперимента, от MEMS датчиков описанным выше способом, которые определяют ориентацию грудного отдела, плечевой кости и локтевой кости. Затем проводится калибровка датчиков на основе алгоритма описанного выше, позволяющая сопоставить движение датчиков с соответствующим телесными сегментами в модели.

Следующим этапом является решение прямой задачи динамики для получения обобщенных сил и моментов, в двух суставах плече и локте. Решение задачи представляет собой интегрирование динамических уравнений модели мускуло-скелетной системы, начиная с заданного начального состояния по описанной схеме 3.4.1. Для решения (численного интегрирования) динамических уравнений, описывающих траектории состояний модели мускуло-скелетной системы на определенном интервале времени, используется интегратор Рунге-Кутты 5-го порядка с адаптивным шагом (Runge-Kutta-Fehlberg).

Далее применяется алгоритм основанный на модели мышц Хила, который позволяет вычислить оценку потребляемой мощности для мышцы тела человека. Мощность метаболизма подсчитывается используя модели [97; 100]. Для того чтобы лучше понять как считается метаболическая мощность мышц далее кратко описывается типы мышц и устройство мышечной ткани.

Мышцы человека классифицируются на несколько типов в зависимости от их структурных и функциональных характеристик. Основные типы мышц включают:

Скелетные (поперечнополосатые) мышцы связаны с движением скелета и контролируют способность к добровольным движениям. Они составляют большинство мышц в организме и часто соединяются с костями через сухожилия. Скелетные мышцы действуют путем сокращения и расслабления в ответ на нервные импульсы.

Гладкие мышцы находятся в стенках внутренних органов, таких как пищеварительный тракт, дыхательные пути и сосуды. Они обеспечивают автоматическое, инволюнтное сокращение органов для поддержания функций, таких как перистальтика, дыхание и регулирование кровообращения.

Мышцы сердца (кардиомиоциты) обладают уникальными свойствами, которые позволяют ей сокращаться ритмично и координированно, обеспечивая кровообращение в организме. Сердечные мышцы образуют миокард и сокращаются автоматически, контролируя ритм сердечных сокращений.

Мышечная ткань, составляющая сократительное компонентное вещество мышц, состоит из множества мышечных волокон, которые могут быть классифицированы на основе их структурных и функциональных характеристик. Существует два основных типа мышечных волокон: медленные волокна (тип I) и быстрые волокна (тип II), которые дополняются подтипами на основе различных свойств.

Медленные (тип I) мышечные волокна характеризуются более высокой способностью к длительной работе без усталости. Они содержат больше митохондрий, что позволяет эффективно использовать кислород для производства энергии путем окисления жирных кислот. Эти волокна обладают более медленной сократительной скоростью, что делает их предпочтительными для активностей, связанных с выносливостью, например, бег на длинные дистанции.

Быстрые (тип II) мышечные волокна имеют большую способность к быстрому сокращению, но обладают более ограниченной выносливостью в сравнении с медленными волокнами. Они подразделяются на два основных подтипа: Тип IIa неутомляемые волокна обладают хорошей способностью к выработке энергии как через аэробные, так и через анаэробные метаболические пути. Эти

волокна используются для среднеинтенсивной деятельности, такой как бег на средние дистанции.

Тип IIb (или IIx) волокна обладают высокой скоростью сокращения и быстрым развитием силы, но быстро утомляются из-за использования анаэробного метаболизма. Они обычно используются для интенсивных короткосрочных активностей, таких как быстрый спринт.

Эти различные типы мышечных волокон обеспечивают человеку адаптивность к различным физическим требованиям, их соотношение в мышцах зависит от генетики, физической активности и тренированности.

За сокращение мышечных волокон отвечает активация мышц, которая представляет собой процесс передачи нервных импульсов от центральной нервной системы к мышцам. Этот процесс начинается с генерации нервных импульсов в моторных нейронах, расположенных в спинном мозге или головном мозге. Нервные импульсы затем распространяются по аксонам нейронов и достигают места контакта нервов с мышцами, называемое нейромышечной синапсом. Этот процесс обеспечивает согласованность и координированность движений организма, а также адаптирует силу и скорость сокращения мышц под различные требования физиологической нагрузки.

В момент достижения мышечных волокон, нервные импульсы вызывают высвобождение химического передатчика ацетилхолина в местах синапсов между моторными нейронами и мышечными волокнами. Это запускает каскад химических реакций что приводит к высвобождению кальция в цитоплазму.

Повышение концентрации кальция в цитоплазме мышечных волокон стимулирует активацию актин-миозиновых мостиков – белковых структур, которые обеспечивают сокращение мышц. При связывании актин и миозин образуется актин-миозиновый комплекс, который использует энергию АТФ для выполнения работы и сокращения мышц.

Следовательно, уровень активации мышц прямо влияет на мощность метаболизма. Более высокий уровень активации требует больше энергии, и, следовательно, приводит к повышению мощности метаболизма.

Уровень активации мышц может изменяться в зависимости от требований движения, нагрузки и физиологического состояния. Путем регулирования частоты нервных импульсов и их силы, нервная система регулирует силу и скорость сокращения мышц, а также координацию движений в организме. Этот

процесс активации мышц является фундаментальным для выполнения физических действий и обеспечивает адаптивную реакцию организма на разнообразные условия и потребности.

Метаболическая мощность мышц (или скорость потребления метаболической энергии) равна скорости выделения тепла плюс скорости выполнения работы:

$$E = B + A + M + S + W$$

где B – базовая скорость выделения тепла (Вт),

A – скорость выделения тепла от активации (Вт),

M – скорость выделения тепла от поддержания (Вт),

S – скорость выделения тепла от сокращения (Вт),

W – скорость механической работы (Вт).

Что бы подсчитать мощность метаболизма мышцы подсчитаем все перечисленные характеристики. Эта модель также использует параметры мышц, согласно мышц работе и справочнику в ней [101]: m – масса мышцы (кг), r – коэффициент медленных волокон в мышце (от 0 до 1) все остальные волокна в мышце основные.

Модель активации (recruitment), описанная Bhargava [100], используется для определения коэффициента медленных волокон, используемого в расчетах ниже. Максимальное значение, используется только при полной возбуждении (т.е. когда все основные волокна активны). С уменьшением возбуждения от 1 до 0 доля активных медленных волокон увеличивается от r до 1.

Базовая скорость выделение тепла (Вт) B представляет собой параметр, который используется для оценки базового уровня энергетического обмена в организме в состоянии покоя. Базовая скорость выделение тепла или Базовый метаболический обмен отражает количество энергии, которое организм расходует для поддержания жизнедеятельности в состоянии абсолютного покоя, когда пища не усваивается и физическая активность минимальна. рассчитывается следующим образом:

$$B = basal_coefficient \cdot (m_{body}^{basal_exponent}),$$

где: *basal_coefficient* – коэффициент базового метаболического обмена. Коэффициент базового метаболического обмена представляет собой множитель, который используется для оценки BMR на основе массы тела организма. Обычно коэффициент базового метаболического обмена включает в себя параметры, такие как возраст, пол и масса тела, m_{body} – масса тела организма, *basal_exponent* – показатель степени, определяющий влияние массы тела на B .

Скорость выделения тепла от активации и поддержания (Вт) $A + M$ рассчитывается следующим образом:

$$A + M = (128(1 - r) + 25) \cdot act^{3/5} \cdot s,$$

при $l_{CE} \leq l_{CE_{opt}}$, и при $l_{CE} > l_{CE_{opt}}$

$$A + M = (0.4 \cdot (128(1 - r) + 25) + 0.6 \cdot ((128(1 - r) + 25) \cdot F_{CE_{iso}})) \cdot act^{3/5} \cdot s$$

$$act = u, \text{ при } u > a$$

$$act = (u + a)/2, \text{ при } u \leq a$$

где m – масса мышцы (кг), l_{CE} – Длина мышечного волокна в текущий момент времени, $l_{CE_{opt}}$ – оптимальная длина волокна мышцы, $F_{CE_{iso}}$ – нормализованная кривая силы-длины сократительного элемента, u – возбуждение мышцы в текущий момент времени, a – активация мышцы в текущий момент времени, s – коэффициент аэробности/анаэробности, определенный свойством движения (обычно 1,0 для преимущественно анаэробных активностей, 1,5 для преимущественно аэробных активностей).

Скорость выделения тепла от сокращения (Вт) S рассчитывается следующим образом при $l_{CE} \leq l_{CE_{opt}}$, $v_{CE} \geq 0$ выражение примет вид:

$$S = m \cdot (-(\alpha_{S_{slow}} \cdot v_{CE_{norm}} \cdot r) - (\alpha_{S_{fast}} \cdot v_{CE_{norm}} \cdot (1 - r))) \cdot act^2 \cdot s,$$

при $l_{CE} > l_{CE_{opt}}$, $v_{CE} \geq 0$ выражение примет вид:

$$S = m \cdot (-(\alpha_{S_{slow}} \cdot v_{CE_{norm}} \cdot r) - (\alpha_{S_{fast}} \cdot v_{CE_{norm}} \cdot (1 - r))) \cdot act^2 \cdot s \cdot F_{iso},$$

при $l_{CE} \leq l_{CE_{opt}}, v_{CE} < 0$ выражение примет вид:

$$S = m \cdot \alpha_L \cdot v_{CE_{norm}} \cdot act \cdot s,$$

при $l_{CE} > l_{CE_{opt}}, v_{CE} < 0$ выражение примет вид:

$$S = m \cdot \alpha_L \cdot v_{CE_{norm}} \cdot act \cdot s \cdot F_{CE_{iso}},$$

где $F_{CE_{iso}}$ – сила, которая была бы развита сократительным элементом мышцы в изометрических условиях при текущей активации и длине волокна, $v_{CE_{max}}$ – максимальная скорость сокращения мышцы, $v_{CE_{norm}}$ – нормализованная скорость сокращения мышечного волокна.

Механическая работа (Вт) W рассчитывается следующим образом:

$$W = -(F_{CE} \cdot v_{CE})$$

где F_{CE} – сила, развиваемая сократительным элементом мышцы в текущий момент времени, v_{CE} – скорость сокращения мышечного волокна. При этом, учитывается первый закон термодинамики, указывающий на необходимость включения отрицательной механической работы в W .

В результате чего получают оценку затраченной энергии человеком E оцененная с помощью мощности метаболизма для последовательности движений человека, на основе данных с микромеханических датчиков. Далее рассматривается алгоритм калибровки датчиков к сегменту.

3.5 Алгоритм калибровки микромеханических датчиков к сегменту

Будем рассматривать движения как движения упрощенной модели человека описанной в разделе 3.2. Предполагается, что каждое из звеньев системы соответствует определенному датчику на теле человека. Ориентация каждого звена определяется по показаниям соответствующего MEMS датчиков. В данной работе рассматривается случай, когда точки крепления всех датчиков к телу человека известны, но ориентация датчика может быть случайной, имен-

но это служит существенным упрощением для работы для человека и отличием для существующих подходов.

Стоит отметить что точка крепления датчика к звену тоже может быть случайной, однако от точки крепления сильно зависит точность восстанавливаемого скелета по нескольким причинам. Самая основная состоит в предположении, что датчики не двигаются относительно костей тела человека. Существует медицинское понятие смещение кожи относительно костей человека называется «подвижностью кожи» или «скольжением кожи». Это явление возникает благодаря тому, что кожа является относительно свободной и гибкой, что позволяет ей перемещаться независимо от подлежащих костей и мышц. И в действительности это допущений может давать ошибку более чем в 15° при ошибочном размещении на тех участках кожного покрова, где подвижность проявляется сильнее всего. Второй причиной является, то что кожа человека эластична и может растягиваться и сжиматься, в таких условиях нельзя сказать, что датчик закрепленный в какой-то точке кожи неподвижен относительно нее, однако этот эффект влияет в меньшей степени и создает ошибку оцененную экспериментально менее чем в 2° . Третьей причиной является, то что мышцы в человеческом теле состоят из множества мелких волокон, которые сокращаются и расслабляются, когда нервные импульсы их стимулируют. Когда мышцы сокращаются, они тянутся и изменяют свою форму тем самым меняют поверхностный рельеф тела человека, что при ошибочном размещении может приводить к ошибке более чем 18° . Однако в ходе экспериментов, была найдена конфигурация взаимного расположение датчиков на теле человека, позволяющая осуществлять устойчивую работу системы сбора данных, схема расположения десяти датчиков на рис. 1.2, однако датчики на предплечьях необходимо расположить на запястьях.

Алгоритм калибровки, не зависит от выбора модели. Поэтому без ограничения общности для простоты представления и описания алгоритма используется упрощенная модель тела человека, состоящая из десяти звеньев: четыре звена моделируют руки человека (плечо и предплечье, без учета кистей рук), другие четыре звена моделируют ноги человека (бедро и голень, без учета стопы) и два звена – позвоночник (плечевая и поясничная области). Каждое звено соединяется с предыдущим сферическим шарниром. Датчики находятся в заранее известных точках на теле человека в соответствии со звеньями модели.

Пример расположения датчиков изображен на рис. 3.4. В действительности данный алгоритм калибровки может быть применен к любой биологической системе кошка, собака, лошадь и др., описанная модель человека взята для упрощения восприятия.

3.5.1 Расположение датчиков на теле человека

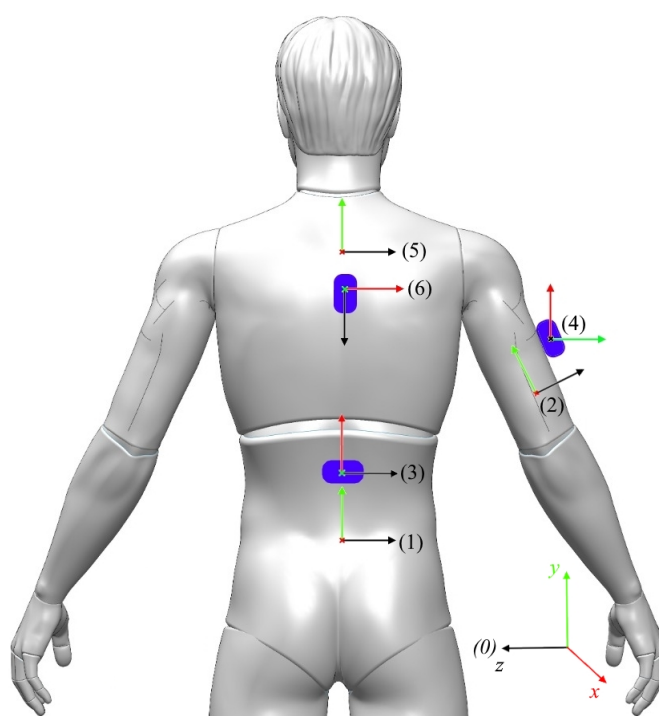


Рисунок 3.4 — Расположение *MEMS* сенсоров (синие) на теле человека.

Цифрами обозначены системы координат:

- (0) – глобальная СК F_{glob} ,
- (1) – СК, связанная с телом человека F_{body} ,
- (2) – СК, связанная с рукой человека F_{arm} ,
- (3) – СК, связанная с датчиком на пояснице человека F_{body_imu} ,
- (4) – СК, связанная с датчиком на руке человека F_{arm_imu} ,
- (5) – СК, связанная с плечевым поясом F_{back} ,
- (6) – СК, связанная с датчиком на плечевом поясе F_{back_imu}

Предложенный алгоритм состоит из двух шагов при которых человек вос-
производит две позы. Без ограничения общности рассматриваются две позы: N

– поза – «руки по швам» – руки вытянуты вдоль туловища от плеч к ступням, при этом плоскость ладони параллельна сагиттальной плоскости человека, и Т-поза – руки разведены в стороны, при этом направление большого пальца коллинеарно направлению взгляда человека «вперед». Заметим, что в общем виде датчики могут находиться на любых частях тела и достаточно того, чтобы части тела друг относительно друга имели объем движения не менее 90 градусов.

3.5.2 Этапы калибровки

На первом этапе калибровки определяется ориентация СК тела в глобальном пространстве. Это реализуется с помощью двух датчиков и данных с двух поз воспроизводимых человеком.

Сначала человек принимает N -позу, и снимаются показания всех датчиков. Показания первого и второго выделенных датчиков задают два вращения: $body_imu R_{glob}(Npose)$ и $arm_imu R_{glob}(Npose)$. Далее человек переходит из N -позы в T -позу, и снимаются показания датчика на руке в T -позе:

$arm_imu R_{glob}(Tpose)$. Два положения руки формируют фронтальную плоскость, соответствующую плоскости, в которой происходило движение руки. Ясно, что сагиттальная плоскость (плоскость симметрии человека) перпендикулярна фронтальной и в ней лежит вектор, параллельный оси вращения руки. Поворот руки из N -позы в T -позу можно выразить через показания датчиков руки:

$$arm_imu_T R_{arm_imu_N} = arm_imu R_{glob}(Tpose) \cdot arm_imu R_{glob}(Npose)^{-1}. \quad (3.16)$$

Осям системы координат тела $body R_{glob}$ соответствуют вектора $\{\vec{x}_b, \vec{y}_b, \vec{z}_b\}$. Представим вращение $arm_imu_T R_{arm_imu_N}$ как кватернион в виде $[w, i, j, k]$,

$$[w, i, j, k] = arm_imu_T Q_{arm_imu_N}, \quad (3.17)$$

Его мнимая часть $[i, j, k]$ будет осью вращения, которая коллинеарна оси \vec{x}_b . Вектор оси вращения необходимо нормировать и представить в глобальной системе

координат:

$$\vec{x}_b = {}^{arm_imu}R_{glob}(Npose)^{-1} \cdot UnitVector \begin{pmatrix} i \\ j \\ k \end{pmatrix}. \quad (3.18)$$

В качестве оси \vec{y}_b выбирается вертикальная ось

$$\vec{y}_b = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad (3.19)$$

но в выкладках используется общий вид: \vec{y}_b .

Заметим, что человек не может сделать поворот руки в строго вертикальной плоскости, вследствие чего оси \vec{x}_b и \vec{y}_b не перпендикулярны. Поэтому необходимо произвести корректировку путем вычитания из вектора \vec{x}_b его вертикальной составляющей и нормирования результата.

$$\vec{x}_b = UnitVector\left(\vec{x}_b - \frac{(\vec{x}_b, \vec{y}_b)}{|\vec{x}_b|} \vec{y}_b\right). \quad (3.20)$$

Ось \vec{z}_b определяется через \vec{x}_b и \vec{y}_b как:

$$\vec{z}_b = \vec{x}_b \times \vec{y}_b. \quad (3.21)$$

Таким образом, можно получить представление СК тела в глобальной СК, составив матрицу вращения из векторов – столбцов $\vec{x}_b, \vec{y}_b, \vec{z}_b$.

$${}^{body}M_{glob} = \begin{pmatrix} \vec{x}_b & \vec{y}_b & \vec{z}_b \end{pmatrix}. \quad (3.22)$$

Для определения положения частей тела по прикрепленным к ним датчикам необходимо знать ориентацию датчиков относительно соответствующих частей тела. Для датчика, закрепленного на пояснице, это вращение рассчитывается по формуле:

$${}^{body}R_{body_imu} = {}^{body}R_{glob} \cdot {}^{body_imu}R_{glob}^{-1}(Npose), \quad (3.23)$$

здесь ${}^{body}R_{glob}$ задается выражением (3.22).

Положение СК частей тела F_{seg} относительно СК тела F_{body} в N -позе ${}^{seg}R_{body}(Npose)$ определяется кинематикой человека, и может быть рассчитано положение всех остальных датчиков используя следующее выражение:

$${}^{body}R_{seg_imu} = {}^{body}R_{glob} \cdot {}^{seg_imu}R_{glob}^{-1}(Npose), \quad (3.24)$$

$${}^{seg}R_{seg_imu} = {}^{seg}R_{body}(Npose) \cdot {}^{body}R_{seg_imu}. \quad (3.25)$$

Заметим, что положение датчиков относительно соответствующих частей тела может быть любым, но фиксированным для каждой калибровки.

Далее определяется текущее положение сегментов модели человека.

После проведения калибровки и нормировки были получены вращения ${}^{body}R_{glob}$ и ${}^{seg}R_{glob}$, соответственно выражения (3.26) и (3.27). Имея полученные данные, можно рассчитать положение частей тела относительно глобальной СК в произвольной позе $Cpose$:

$${}^{body}R_{glob}(Cpose) = {}^{body}R_{body_imu} \cdot {}^{body_imu}R_{glob}(Cpose). \quad (3.26)$$

$${}^{seg}R_{glob}(Cpose) = {}^{seg}R_{seg_imu} \cdot {}^{seg_imu}R_{glob}(Cpose). \quad (3.27)$$

Положение частей тела относительно СК тела рассчитывается следующим образом:

$${}^{seg}R_{body}(Cpose) = {}^{seg}R_{glob}(Cpose) \cdot {}^{body}R_{glob}^{-1}(Cpose). \quad (3.28)$$

Рассмотренный алгоритм калибровки был опробован в системе сбора и визуализации данных, описанной ниже.

3.6 Практическое применение системы сбора данных

Архитектура системы сбора данных состоит из компьютера под ОС *Windows*, набора из десяти MEMS датчиков, роутера и системы визуализации. Набор датчиков в своей основе использует промышленные девятиосевые IMU сенсоры *BNO055* от компании *Bosch*. Управление данными с датчиков выполняет двухъядерный микроконтроллер *ESP32* с низким энергопотреблением с интегрированным *Wi-Fi* и *Bluetooth* контроллерами и антеннами. Система мик-

роконтроллер и датчики работают автономно от *Li-ion* аккумулятора емкостью 800mA. Данного аккумулятора хватает на 8 часов непрерывной работы датчиков. На данном этапе исследований данные с датчиков собираются контроллером и отправляются по *UDP* протоколу по *WI-Fi* сети через роутер на компьютер. Перед сбором данных проходит процесс синхронизации времен между каждым датчиком и компьютером, при котором считается время прохождения сигнала между ними.

3.6.1 Система визуализации данных

Система визуализации написана на языке *Python* с использованием библиотек *PyQT5* – (кроссплатформенный фреймворк) и *VTK* – (открытая кроссплатформенная библиотека для трёхмерного моделирования, обработки изображений и прикладной визуализации). В системе визуализации за основу принята система координат в *OpenGL* (спецификация, определяющая платформо-независимый программный интерфейс для написания приложений, использующих двумерную и трёхмерную компьютерную графику). В глобальной системе координат, по умолчанию используемой в *BNO055*, ось \vec{z} направлена вверх, противоположно вектору гравитации, а ось \vec{y} направлена на север. Такая же система координат используется во многих других MEMS датчиках. В системе визуализации удобнее работать в СК, применяемой в *OpenGL*: ось \vec{y} направлена вверх, а ось \vec{z} направлена на камеру, поэтому мы переводим данные приходящие с датчиков в СК *OpenGL*, используя следующую матрицу поворота:

$${}^{OpenGL}M_{BNO055} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}. \quad (3.29)$$

В системе визуализации используются два режима отображения: в СК, связанной с телом человека, когда фигура все время повернута лицом к камере, и в глобальной СК, когда поворот тела отображается на экране. Для второго варианта отображения дополнительно к процессу калибровки можно ввести

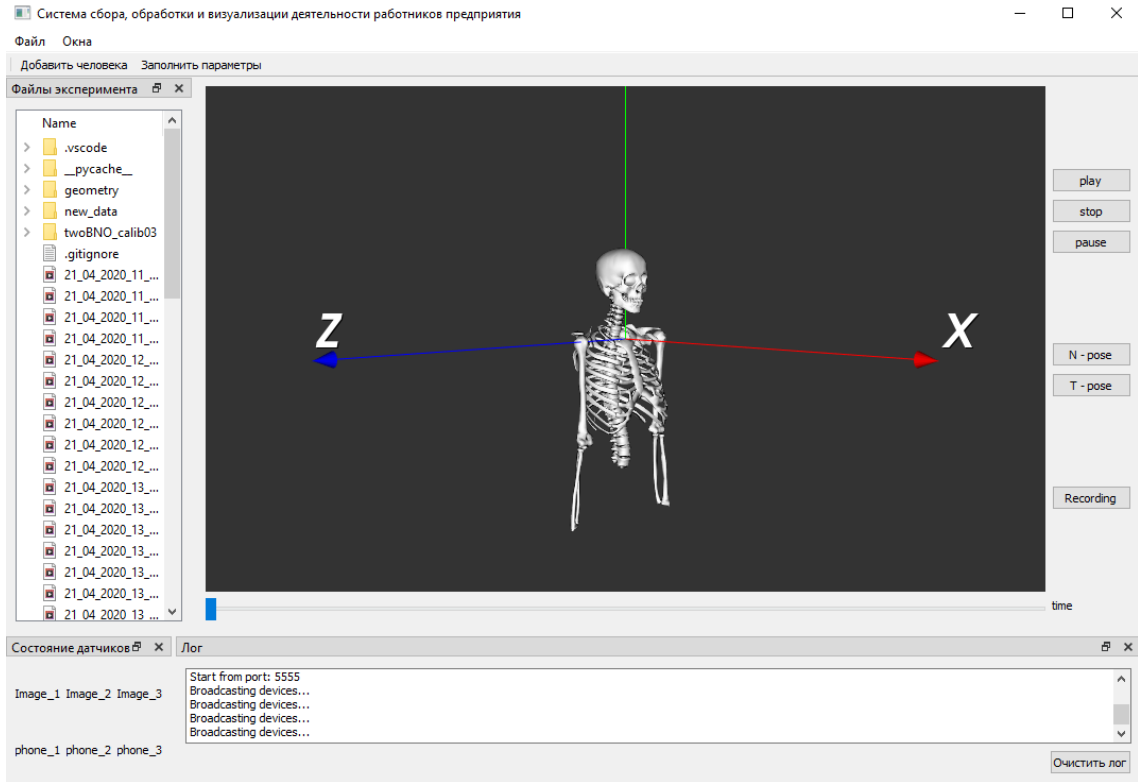


Рисунок 3.5 — Внешний вид системы визуализации

процесс нормировки, который приводит откалиброванное положение модели в глобальной модельной системе координат F_{scene_glob} к начальному.

После проведения калибровки и нормировки были получены вращения ${}^{body}R_{glob}$ и ${}^{seg}R_{glob}$, соответственно выражения (3.26) и (3.27).

$${}^{scene_body}R_{glob}(Cpose) = {}^{body}R_{glob}(Cpose) \cdot {}^{body}R_{glob}(Npose)^{-1}, \quad (3.30)$$

$${}^{scene_seg}R_{glob}(Cpose) = {}^{seg}R_{glob}(Cpose) \cdot {}^{seg}R_{glob}(Npose)^{-1}. \quad (3.31)$$

Таким образом, результаты выражений 3.30 поступают как входные данные ориентаций систем координат тела и всех его сегментов. Внешний вид приложения представлен на рисунке 3.5, а открытый код доступен на сайте [102]. Приложение обрабатывает данные, получаемые с датчиков параллельными процессами, синхронизирует их и отображает на модели.

Демонстрация работы откалиброванной программы приведена на рисунке 3.6.

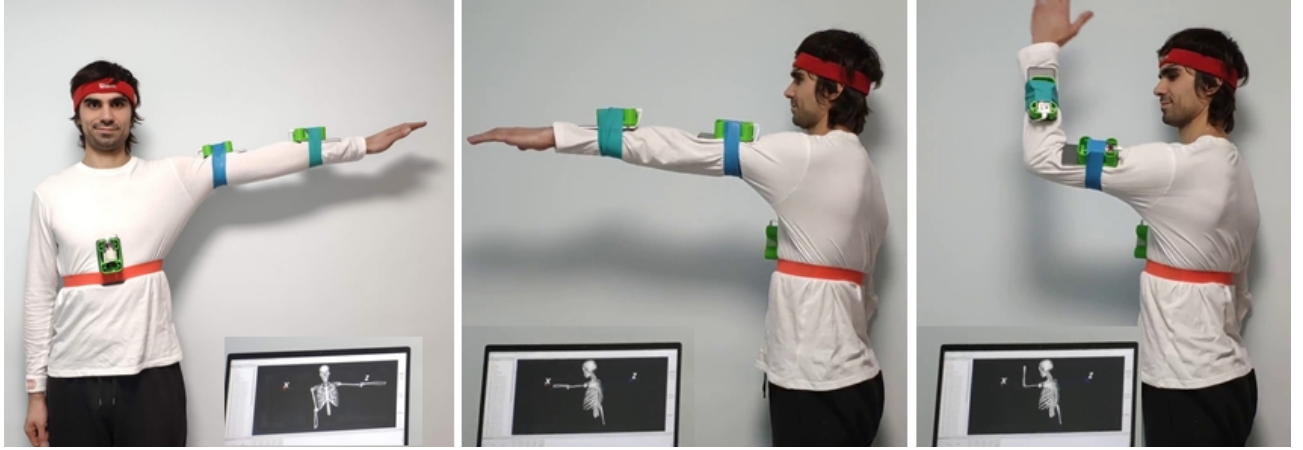


Рисунок 3.6 — Иллюстрация работы калибровки и визуализации

Данный подход предоставляет возможность рационализировать процедуру калибровки, придавая использованию датчиков более эффективный и эргономичный характер, а также уменьшая затраты времени на побочные операции.

3.6.2 Предобработка кинематических данных в формате R3DG

Если рассмотреть формат *R3DG* описанный в работе [90] подробнее, то для каждого из звеньев модели был получен объект *Special Euclidean Group* $SE(3)$ представляющий собой шестимерную матричную группу Ли, образованную набором всех матриц 4×4 вида

$$P(R, \vec{d}) = \begin{bmatrix} R & \vec{d} \\ 0 & 1 \end{bmatrix}, (\vec{d}) \in \mathbb{R}^3, R \in SO(3).$$

Алгебра Ли $SE(3)$, обозначаемая $se(3)$, представляет собой шестимерное векторное пространство, натянутое на набор всех косых матриц 4×4 .

$$\log_{SE(3)}(P) = \log(P).$$

Логарифмическая карта не уникальна в случае $SE(3)$. В этой работе используется $\log(P)$ с наименьшей нормой. Учитывая относительную геометрию между всеми парами частей тела, формируется модель C в момент времени t , используя:

$$C(t) = P_{1,2}(t), P_{2,1}(t), \dots, P_{M,M-1}(t),$$

где представление $P_{i,j}(t) = P_i(t)P_j(t)^{-1}$, означает, что переход одной системы координат в другую представлен в виде сдвига и поворота в системе координат последней.

Каждое движение человека рассматривалось на отрезке $[0; T']$. Используя предложенное представление, скелетную последовательность, описывающую действие, можно представить в виде кривой $[0; T_0]$, а распознавание действий может быть выполнено путем классификации таких кривых по разным категориям действий.

Здесь кривая действия $C(t)$ эволюционирует в неевклидовом пространстве $SE(3)^K$ как функция времени, и стандартный анализ Фурье для этого случая не определен. Для преодоления этих трудностей составим карту кривых действия из группы Ли $SE(3)^K$ в ее алгебру Ли $se(3)^K$, которая представляет собой $6M(M-1)$ -мерное векторное пространство. Окончательное представление кривой действия $C(t)$ имеет вид

$$c(t) = [\text{vec}(\log(P_{1,2}(t))), \text{vec}(\log(P_{2,1}(t))), \dots, \text{vec}(\log(P_{M-1,M}(t))), \\ \text{vec}(\log(P_{M,M-1}(t)))]$$

3.6.3 Предобработка кинематических данных в формате kinect (.skeleton)

Про практическое применение *Microsoft Kinect V2* написано в работе [103]. Формат данных *.skeleton* в данной работе можно получить как из данных видео системы так и из данных на микромеханических датчиков так как и тот и тот алгоритмы применяются к одной и той же компьютерной модели для отображения. В разделе будет приведены вычисления на основе данных *MEMS* – датчиков, но это не ограничивает применение и для системы на основе видео данных. По информации с *MEMS* датчиков и известным длинам костей воссоздается кинематика движения человеческого тела. В рамках восстановления кинематики введена глобальная система координат модели. Нам известна ориентация системы координат человеческого тела, относительно глобальной системы координат модели ${}^{body}F_{globe}$.

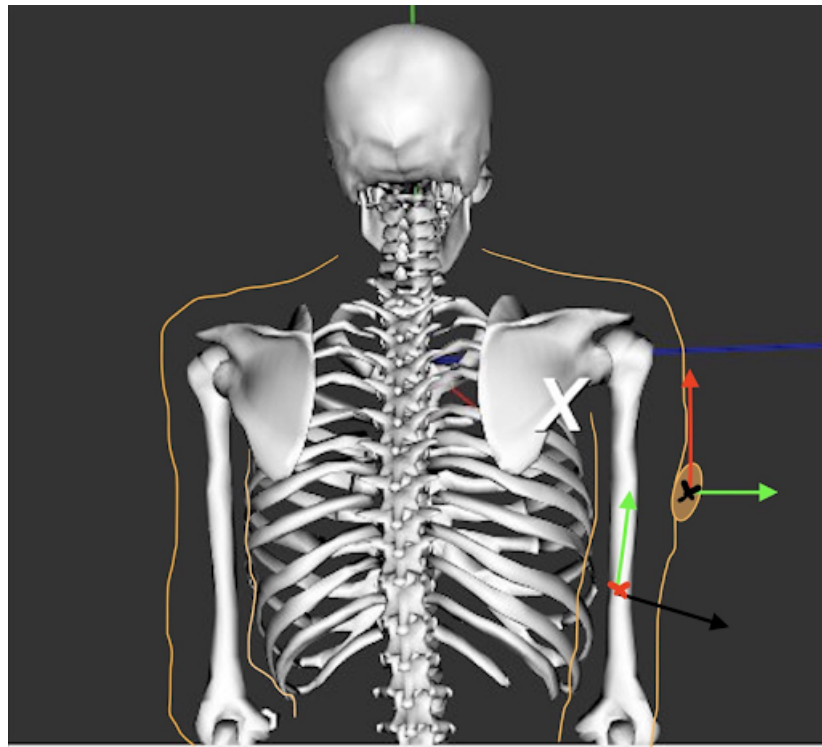


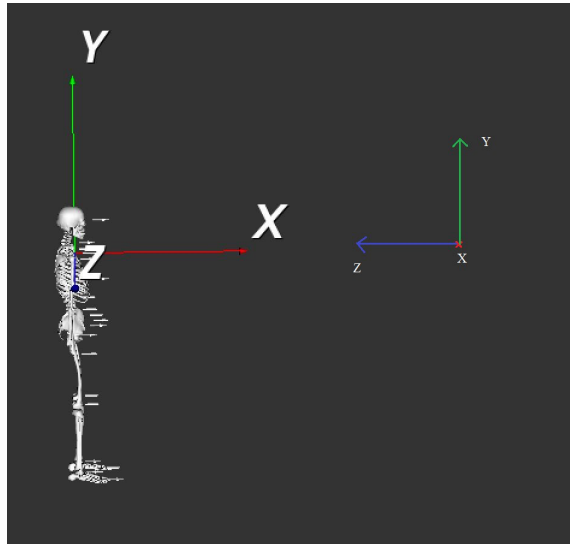
Рисунок 3.7 — СК сегмента и СК MEMS сенсора.

Также известно вращение сегмента, на котором расположен датчик (рис. 3.7), относительно системы координат тела. Компоненты $^{scene_body}F_{body_imu}$ и $^{scene_seg}F_{seg_imu}$ получены из калибровки, а $^{body_imu}F_{glob}$ и $^{seg_imu}F_{glob}$ измеряются с помощью MEMS датчиков.

В глобальной системе координат задана система координат $kinect K_{globe}^{\rightarrow}$ с центром в некоторой точке, которая рассчитывается как угол прямоугольника в который вписана траектория будущего движения человека по двум координатам X, Z . Также задан кватернион системы координат $kinect$ относительно глобальной системы координат определяющий направление $kinect$ в точку пересечения диагоналей упомянутого прямоугольника $^{scene}K_{glob}$.

Для определения ориентации сегмента в системе координат $kinect$. Необходимо преобразовать вращения сегментов из глобальной системы координат в систему координат $kinect$.

Из координат СК $kinect$ рис 3.8 в глобальной системе координат вычитаем координаты сустава в глобальной СК. Полученный вектор поворачивается на сопряженный кватернион системы координат $kinect$ относительно глобальной системы координат.

Рисунок 3.8 — СК тела человека и СК *kinect*.

$$\vec{F}_{scene_seg} = K_{globe} \vec{F}_{globe_seg}, \quad (3.32)$$

$${}^{scene_seg} F_{globe} = ({}^{scene_seg} F_{body_imu} \cdot {}^{body_imu} F_{glob}) \cdot {}^{scene} K_{glob}. \quad (3.33)$$

Полученные координаты сегмента в системе координат *kinect* соответствуют координатам относительно камеры *kinect* 3.32. При этом кватернион сегмента соответствует кватерниону, связанному с суставом в формате *kinect* 3.33.

Для того чтобы заполнить все данные в соответствии с форматом *kinect*, требуется вычислить координаты цвета и координаты глубины. Координаты цвета соответствуют положению центральной точки сустава на *RGB* изображении, а координаты глубины соответствуют положению центральной точке сустава на *IR* изображении.

В качестве входных данных есть координаты сегмента в системе координат *kinect*, угол обзора камеры, ширина и высота экрана, на который происходит проекция. Для того чтобы вычислить искомые данные, нужно вычислить фокусное расстояние камеры. Угол обзора *kinect* $\phi = 43$ градусам. Фокусное расстояние f считается из геометрии, изображенной на рисунке 3.9, по формуле:

$$f = \frac{h}{2 \cdot \text{th}(\frac{\phi}{2})},$$

где высота экрана $h = y_{max} - y_{min}$ вычисляется из координат всей траектории движения. После рассчитываются координаты относительно СК расположен-

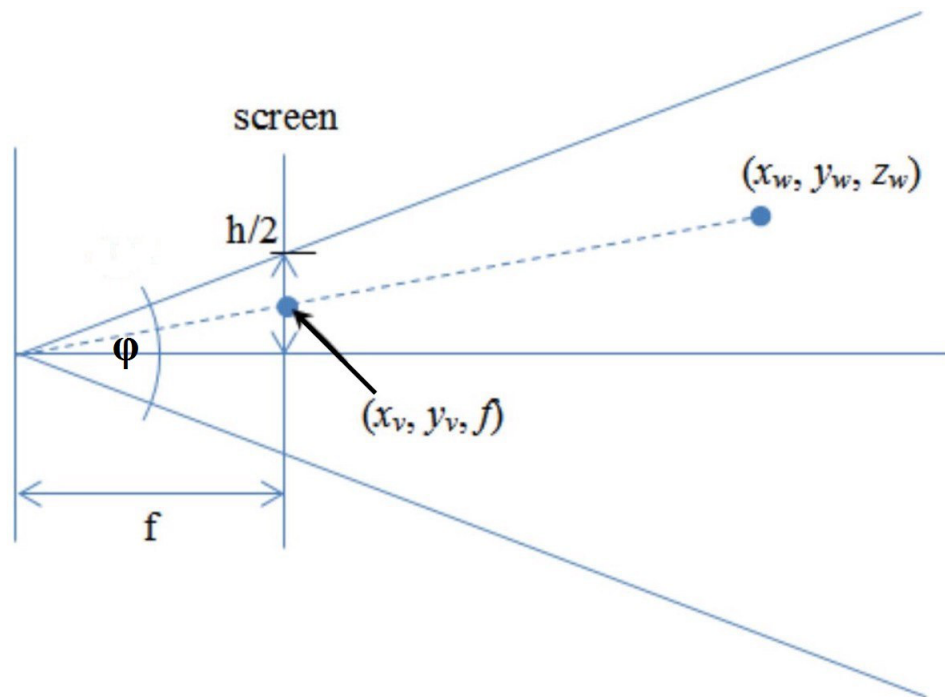


Рисунок 3.9 — Проекция на экран

ной в середине мнимого экрана *kinect* на который происходит проекция. Далее координаты переводятся в систему координат экрана с началом координат в левом верхнем углу и нормируются таким образом, чтобы получить значение в пределах от 0 до 1, как в формате *kinect*.

3.7 Результаты

В результате описанных алгоритмов, была собрана программно аппаратная система состоящая из:

1. Системы сбора кинематических данных (датчики, браслеты, головное устройство (ноутбук));
2. Системы камер (от 1 до 4х);
3. ПО для сбора синхронизации и фильтрации всех получаемых данных;
4. ПО калибровки и восстановления трехмерных моделей человека по датчикам или камерам.

На основе системы было собрано модуль который, принимает на вход кинематические данные от датчиков расположенных на теле человека и восстанавливает трехмерную модель тела человека в реальном времени. Такой же мо-

дуль был создан для видео системы. Каждый модуль обладает возможностью выгрузки обучающих наборов данных в известных форматах (*.myn*, *.skeleton*).

Глава 4. Анализ последовательностей движений человека на основе совокупности видео данных и данных микроэлектромеханических датчиков

4.1 Классификация последовательностей движений человека

в современной практике машинного обучения сбор большого объема данных стал важной составляющей успешного обучения моделей, особенно в случае использования сложных и глубоких алгоритмов.

1. Увеличение сложности задач: Современные алгоритмы машинного обучения, такие как глубокие нейронные сети, способны решать более сложные задачи, чем простые линейные модели. Для обучения таких сложных моделей требуется больше данных, чтобы модель смогла извлечь значимые закономерности из информации.
2. Улучшение качества прогнозов: Чем больше данных доступно для обучения модели, тем точнее и надежнее будут ее прогнозы. Большие объемы данных позволяют модели лучше учиться и делать более точные предсказания.
3. Избегание переобучения: Когда модель сложна, а данных недостаточно, существует риск переобучения, когда модель «запоминает» обучающие данные, но не способна обобщать на новые данные. Добавление большего объема данных может помочь уменьшить риск переобучения и сделать модель более обобщающей.
4. Разнообразие данных: Разнообразие данных имеет важное значение для успешного обучения моделей машинного обучения. Чем больше данных разных типов и из разных источников, тем лучше модель способна учить обобщенные закономерности.
5. Улучшение обучения на больших наборах данных: Современные алгоритмы машинного обучения, особенно глубокие нейронные сети, показывают выдающиеся результаты на больших наборах данных. Это связано с тем, что большие объемы данных позволяют использовать масштабируемые вычисления и более точно настраивать параметры модели. Масштабируемость вычислений – это концепция и подход в

информационных технологиях, который связан с способностью системы или алгоритма эффективно обрабатывать увеличивающееся количество данных или нагрузку при добавлении вычислительных ресурсов. Масштабируемость – это способность системы или программы поддерживать свою производительность и функциональность при увеличении объема работы или данных.

6. Обучение без учителя: Даже алгоритмы машинного обучения без учителя, такие как кластеризация и вариационные автокодировщики, могут потребовать больших объемов данных для выявления структуры в данных или создания хороших представлений.

При этом для применения алгоритма в реальных условиях с точки зрения экономической целесообразности и надежности распознавания желателен минимальный набор датчиков. Стоит также сказать, что действительно впечатляющих результатов удастся достичь при учете специфики задачи в алгоритме. В данном разделе описаны алгоритм классификации последовательностей движений (действий) человека, сбор датасета с инерциальных датчиков о движении человека в лабораторных условиях с видео, универсальные форматы данных о движении и модели классификации, обученные на различных форматах.

В данной задаче исследуется проблема классификации последовательностей движений человека, выполняющего работы с инструментом. Будет рассмотрено 3 задачи с разными входными данными, но с одними целями. Задача классификации последовательностей движений:

1. на основе кинематических данных с датчиков в локальной системе координат человека полученных в разделе [2.1](#).
2. на основе кинематических данных с датчиков преобразованных в многомерные кривые по методу R3DG полученных в разделе [3.6.2](#)
3. на основе кинематических данных с датчиков преобразованных в формат *.skeleton* полученных в разделе [3.6.3](#).
4. на основе видео данных с распознаванием ключевых точек человека алгоритмом *OpenPose* полученных в разделе [1.3.1](#).
5. на основе видео данных с распознаванием ключевых точек человека алгоритмом *OpenPose* и восстановлением 3D позы по данным с двух камер полученных в разделе [1.4](#).

6. на основе видео данных с распознаванием ключевых точек человека алгоритмом *OpenPose*, восстановлением 3D позы и конвертированных в формат *.skeleton*.

Итоговое наиболее робастное решение для поставленных целей подробно будет описано в данном разделе. Все результаты полученные в ходе исследования приведены в заключении к этой главе. Для классификации используются данные инерциальных датчиков, видео данные, – собираемые в эксперименте, участвуют только в разметке данных с инерциальных датчиков. Исследование позволило получить следующее решение: собрать данные группой инерциальных датчиков; обработать их с помощью математической модели человека; интерпретировать движение модели в формате *.kinect (skeletal data)*, собрать данные группой видеокамер, разметить видеокадры, перенести разметку на данные с инерциальных сенсоров. На основе собранных данных сформировать собственные обучающие выборки. Далее, пользуясь результатами обзора датасетов приведенного во введении, собрать из открытых датасетов максимально большой датасет со схожими действиями для обучения в том же формате. Создать набор разработанных нами признаков. На основе собранных выборок обучить модель классификации один против остальных (OVR) описанный в 2.1, для мультиклассового обучения, на основе метода *xgboost* [104]. Далее рассмотреть временную последовательность во время совершения действия и убрать шумы при работе модели для использования ее в реальных условиях, а также для улучшения предсказания моментов начала и конца действий из последующей задачи [86].

4.1.1 Целевые переменные

Для каждого видео была создана таблица, в которой указывался номер кадра и три бинарные целевые переменные, соответствующие этому кадру, которые нужно предсказывать в задаче. Первая переменная — поднята ли правая рука у человека на видео выше 90° относительно туловища, где 1 – угол между плечом и туловищем человека больше или равен 90° , 0 – иначе. Вторая переменная – красит ли стену человек на видео, где 2 – человек красит стену

(валик соприкасается со стеной), 0 – иначе. Третья целевая переменная определяет иное, но конкретное действие – 3, и является 0 в остальных случаях включая переходные действия или моменты ожидания. Заполнение значений целевых переменных проводилось вручную путем покадрового просмотра полученных данных с видео. Далее видеокадрам сопоставлялись – по временной сетке записи инерциальных датчиков. Таким способом, осуществлена разметка данных о положении человека.

4.1.2 Формирование общего набора данных

Рассматриваются четыре открытых датасета (UTD – MHAD [105], UT – Kinect [56], NTURGB + D [106]), из каждого выбираются действия, удовлетворяющие первой (поднята ли правая рука у человека) и третьей (конкретное иное действие) целевой переменной. В итоге были взяты действия из датасетов UTD – MHAD (1, 2, 3, 5, 7, 8, 10, 17, 19, 20), UT – Kinect (2, 4, 5, 15), MSRC – Kinect12 (4, 7, 8), NTURGB + D (4, 7, 18, 19, 20, 21, 28, 38, 44, 45, 61, 62, 104), номерами обозначены взятые последовательности движений–действия в качестве первой целевой переменной, оставшиеся действия – в качестве третьей. В итоговом векторе собранных признаков основанном на данных в формате *.skeleton* 3.6.3 присутствует максимальное количество распознанных ключевых точек (суставов) скелета для данной выборке это число суставов в наборе с наименьшим количеством и это 12 суставов. В кадрах, где присутствуют два человека, каждый человек был рассмотрен в отдельности.

4.1.3 Получение дополнительных признаков

Все подготовленные данные приводились к единой временной шкале в зависимости от частоты сбора данных, и время выступало одним из признаков для классификации в качестве монотонно возрастающей последовательности. Дата была разделена на: дни недели, дни в месяце, время дня, для некоторых

выборок это не будет иметь результата, но в своей общности на основе опыта проведения промышленных испытаний было сделано предположение, что существует корреляция между работой на предприятии и временем. Общее описание: в своей основе данные без пропусков, количественные, 25 частей тела человека с 11 свойствами, главными из которых являются координаты фиксированной точки части тела, кватернионы (циклические комплексные переменные, задающие ориентацию части) 3.3.2 и глубина. Глубина и координаты являются линейными характеристиками, которые медленно изменяются и просты в анализе. Кватернионы – нелинейные характеристики, с выбросами, корреляций с целевыми переменными отсутствуют. В общности перечисленных признаков мало для построения классификатора так как отсутствуют корреляции с целевыми переменными при значительном дисбалансе классов. Глубина и координаты для каждой части тела были переведены в локальную систему координат, связанную с человеком, и нормированы на длину от шеи до центра бедер для координат, и карты глубин человека отделили от фона.

Для исключения выбросов в координатах были апробированы разные методы, в том числе метод межквартильного диапазона [107] и разные меры в том числе и мера Махаланобиса [108]. В результирующей выборке выбросы и примеры с ложным распознаванием были исключены фильтрацией по стандартному отклонению значений пространственных координат, то есть были оставлены лишь те, что попадали в 95% доверительный интервал по пространственным координатам для каждого набора данных в отдельности. А так же дополнительно все данные были обработаны системой контроля скелетов 1.3.5.

Введем предположение, основанное на наблюдении человеческих действий на производстве: действия человека относительно изменений кватернионов во времени является низкочастотной характеристикой. Вот несколько причин, почему это может быть так:

Физические ограничения человеческого тела: Физические ограничения, такие как инерция и анатомическая структура тела, могут ограничивать скорость изменения ориентации. Например, чтобы выполнить поворот или сменить позу, человеческое тело должно преодолеть инерцию и учитывать свою анатомию, что ограничивает скорость изменений.

Биомеханика движений: Движения человека, такие как ходьба или бег, обычно имеют определенные биомеханические характеристики, которые опре-

деляют скорость изменений. Эти характеристики могут быть связаны с физиологическими ограничениями и эффективностью движений, а так же привычкой точностью повторения одних и тех же действий раз за разом.

Человеческие восприятия и управление: Человек воспринимает и управляет своими движениями на основе сенсорной обратной связи и целей задачи. Медленные и плавные изменения ориентации могут быть более управляемыми и предсказуемыми для человека.

В том числе и из-за шумов, порождаемых технической реализацией датчиков. На основе опытных испытаний целевые действия длятся на протяжении 0,5-1-2 секунд. На основе данного предположения будет введено несколько признаков.

С помощью разложения выборки кватернионов в дискретный ряд Фурье отбираются 3 самые представленные частоты в спектре от 0.6 до 0.9 Гц. Так как действия человека могут быть отражены не только в частотной области, но и во временной области, координаты кватерниона обрабатываются вейвлет-преобразованием, которое имеет хорошую частотно-пространственную локализацию. Очевидно, что в сигнале присутствуют локализованные разномасштабные процессы, можно предположить, что рассматриваемое семейство действий будет иметь один локализованный масштаб. Следовательно, будем рассматривать только этот масштабный уровень разложения и кратные ему. В качестве базисной функции берется МНАТ-вейвлет [109], вейвлет симметричен, спектр вейвлета представлен только действительной частью и хорошо локализован по частоте, часто используется для сложных сигналов [109]. После проведения статистического анализа данных была показана возможность утверждать о нормальности распределений всех наблюдаемых признаков. Линейные корреляции в итоговых данных с целевыми переменными отсутствуют. Однако очевидно, что первую целевую переменную угол между плечем и туловищем можно посчитать явно на основе двух кватернионов двух объектов бедра и плеча. Тем не менее, намерено такой признак не будет вводиться для последующей минимизации исходных данных, а именно исключения кватерниона отвечающего за положение плеча, однако эту информацию косвенно можно получить из положения предплечья.

4.1.4 Подготовка обучающей выборки

Итоговый датасет получается методом даунсемплинга [110] – исключением некоторых преобладающих данных в описанной выборке. Выбор модели обусловлен вектором признаков имеющим размерность $11+12(\text{fft})+4*10(\text{cwt})+3(\text{t})$, где fft (Fast Fourier transforms) – конкатенация векторов разложенных в ряд фурье по упомянутым выше частотам, cwt (Continuous wavelet transform) – конкатенация векторов после вейвлет преобразования. Так как в итоговых данных было более 10 тысяч положительных примеров на каждый класс, в качестве оценки будет использоваться кросс-валидация: разбиение на 4 подвыборки для каждого класса, для лучшей интерпретируемости результатов и снижения эффекта переобучения.

Интерпретируемость результатов:

Улучшенная статистическая значимость: Когда вы разбиваете данные на множество подвыборок и проводите кросс-валидацию, результаты анализа становятся более статистически значимыми. Это означает, что вы можете быть более уверенными в том, что обнаруженные закономерности или различия действительно существуют в данных, а не случайны.

Стабильность выводов: Использование нескольких разбиений позволяет проверить, насколько стабильны результаты. Если результаты модели сильно изменяются при разных разбиениях, это может быть признаком нестабильности модели или могло бы свидетельствовать о наличии выбросов в данных.

Оценка обобщающей способности: Кросс-валидация помогает оценить, насколько хорошо модель будет работать на новых, ранее не виденных данных. Это важно для интерпретации результатов, так как позволяет понять, как модель будет вести себя в реальных условиях.

Снижение эффекта переобучения: Снижение переобучения: Разбиение данных на множество подвыборок и проведение кросс-валидации помогает снизить риск переобучения модели. Если бы вы обучали модель на всем наборе данных, существовала бы опасность, что она могла бы «запомнить» данные, вместо того чтобы обобщать информацию.

Более объективная оценка качества модели: Кросс-валидация предоставляет более объективную оценку качества модели, так как она учитывает ее про-

изводительность на разных частях данных. Это позволяет более точно оценить, насколько хорошо модель справляется с задачей классификации, минимизируя влияние случайных особенностей в данных.

В финальной версии дата сета были исключены признаки с разложением в дискретный ряд Фурье по причине нехватки вычислительных ресурсов, оставшиеся описанные атрибуты остались. Распределение на выборки было со стратификацией по целевой переменной для каждого класса.

4.1.5 Результаты обучения

Для мультиклассовой классификации был использован подход один против всех 2.1, что позволило свести задачу к бинарной классификации, так как Все эксперименты по обучению проводились на наборах данных с балансированными классами, которые были получены методом исключения. Окончательная точность модели была оценена на основе средних значений элементов матрицы ошибок (*True Positives–TP*, *False Positives–FP*, *True Negatives–TN*, *False Negatives–FN*) между всеми бинарными классификаторами. Затем по одной усреднённой матрице ошибок считались метрики *precision*, *recall*, *f*-мера [111].

Для демонстрации результатов обучения для кросс-валидации были выбраны метрики также *roc-auc*, *precision*, *recall*, *f1* в таблице 9. Приведен пример полученных результатов по упомянутым метрикам для второй целевой переменной – распознавание действия “красит ли стену человек” (касается ли кистью стены). На основе описанного подхода были обучены 2*5 моделей для усреднения результатов обучения, каждого из 5 случайных разбиений исходного набора данных для кросс-валидации. Итого каждая модель имела 50 вариаций, из которых была выбрана лучшая и впоследствии эта модель была дообучена на полной выборке. Далее приведены точности работы модели на трех классах. Метрика *roc-auc* подсчитана для мультиклассовой классификации по макро усреднению [112].

Полученные результаты улучшаются, на основе методов подбора макропараметров, основные методы подбора макропараметров это сеточный поиск – полный перебор, случайный поиск, байесовской оптимизации – использует-

Таблица 9 — Метрики оценки модели классификации XGBoost, для каждой подвыборки (1-4)

№	<i>Roc-auc</i>	<i>Precision</i>	<i>Recall</i>	<i>f1</i>
1	0.986571	0.929416	0.954931	0.942001
2	0.986652	0.930264	0.953842	0.941905
3	0.986287	0.927319	0.956930	0.941892
4	0.986180	0.927754	0.956270	0.941796

ся статистическая модель, которая аппроксимирует зависимость между макропараметрами и производительностью модели затем применяется алгоритм оптимизации, который исследует пространство параметров и находит наилучшую комбинацию, адаптивная оптимизация параметров – на основе алгоритма *Tree – structuredParzenEstimator(TPE)* [113] для поиска оптимальных значений параметров, он основан на итеративной оценке апостериорного распределения, чтобы найти наилучшую комбинацию параметров и последний метод это гибрид двух последних *SequentialModel – BasedOptimization(SMBO)* [114]. В каждом из методов необходимо задать метрику оценки качества модели, классической метрикой для этих целей является метрика *f1*. Для подбора параметров были апробированы методы сеточного поиска, байесовской оптимизации и (*SMBO*), оптимальный результат по быстродействию и точности показал метод поиска по сетке.

Для дальнейшей оценки было проведено исследование значимости признаков в поставленной задаче классификации на основе работы [115]. Считается важность перестановки для оценки признаков. Важность перестановки признака рассчитывается в два этапа. В первом этапе базовая метрика, определяемая оценкой, тестируется на (потенциально другом) наборе данных, определяемом исходными данными для обучения. Затем столбец функций из набора проверки переставляется, и метрика оценивается снова. Важность перестановки определяется как разница между базовой метрикой и метрикой перестановки столбца признаков, что дает статистическую оценку значимости признаков для классификатора. В результате признаки были упорядочены от наиболее важных к наименее важным: время в минутах от начала дня, кватернионы, их вейвлет преобразования и координаты.

Метрика для мультиклассовой классификации `macro roc-auc` показала точность в 98.8% 4.1. Чтобы достичь «`macro ROC-AUC`», модель должна иметь очень высокий уровень разделения между классами и показывать отличные характеристики классификации для всех классов в среднем. Это означает, что кривые ROC для каждого класса имеют значительную площадь под ними и приближаются к идеальному случаю, где верно классифицированы все положительные и отрицательные случаи.

На графике, который используется для визуализации ROC-кривых, кривая для «`macro ROC-AUC`» в 98.8% будет находиться в верхней левой части графика, ближе к верхнему левому углу, что указывает на высокий уровень производительности модели.

График ROC-AUC обычно строится путем варьирования порога для классификации и измерения чувствительности (Recall) и специфичности (Specificity) модели на разных порогах. Метрика «`macro ROC-AUC`» учитывает среднее значение AUC для всех классов, а не учитывает дисбаланс классов.

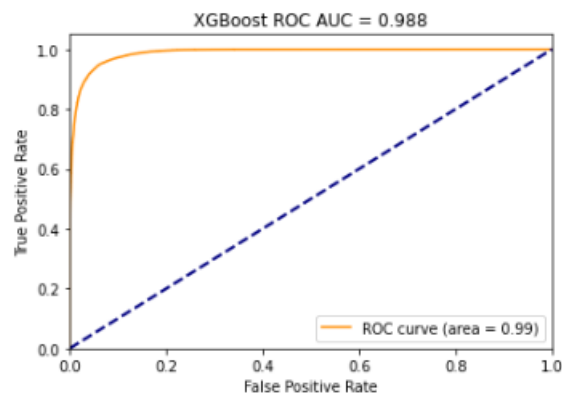


Рисунок 4.1 — Точность по метрике `macro roc-auc`

4.1.6 Вычислительная сложность и оценка времени работы

При сравнения полной цепочки процессов для оценки сложности и времени работы алгоритмов становится очевидно, что вычислительные мощности по обработки изображений с помощью системы OpenPose требуют значительные вычислительные мощности, в отличии от модели на MEMS данных. В данном разделе не представлены численные характеристики по времени работы, так

как система в реальном времени находится в разработке. Отдельным вопросом в данной тематике служит как сравнивать алгоритмы работающие на разного рода процессорах: OpenPose + Classifier работающие как на GPU так и на CPU, и Classifier работающий только на CPU, можно запустить алгоритм OpenPose на CPU, но смысла в анти-оптимизации для сравнения нет. Стоит отметить что вычислительные мощности требующиеся для работы двумя описанными подходами не сравнимы, так как для обработки видео последовательностей необходима система технического зрения с значительными вычислительными мощностями, а для работы датчиков необходима MEMS – система одевающаяся на человека, и требующая калибровки при каждом использовании.

Переходя к практической применимости хотелось отметить что основной проблемой в промышленном тестировании была шумность результатов работы модели, так как частота получения в одну секунду данных эксперимента составляет 100 Гц. При средней точности модели в 98% от одного до трех показаний в секунду ложны. Для решения этой проблемы над последними ста измерениями проводится дообработка данных для большей достоверности результатов и в том числе для более точного определения и взаимодействия с системами из последующих задач [86]. Был применен метод *DBSCAN* (*Density-Based Spatial Clustering of Applications with Noise*) – это алгоритм кластеризации данных, который использует плотность данных для определения кластеров. Он ищет плотные области в пространстве данных и объединяет их в кластеры.

Алгоритм начинается с выбора наиболее достоверной точки в данных и поиска всех соседних точек, которые находятся на расстоянии не более, чем заданный радиус десять измерения. Если вокруг точки есть более, чем минимальное количество соседей в данной работе двумя, то эта точка становится частью кластера. Подобранные величины обусловлены тем, что даже самые быстрые действия человека не могут прерваться на значения в доли секунды, а затем продолжаться.

Затем алгоритм ищет соседние точки этой точки и добавляет их в кластер, если они удовлетворяют критерию плотности. Процесс продолжается до тех пор, пока все точки не будут просмотрены.

4.1.7 Заключение

Сравнение методов видеофиксации без использования и с использованием данных микромеханических датчиков проводилось в лабораторных и промышленных испытаниях данной системы и обучены модели классификации действий человека в условиях промышленного производства. На основе анализа важности признаков была проведена минимизация входного вектора удалось минимизировать количество датчиков до двух. При использовании метода *xgboost* для мультиклассовой классификации удалось добиться точности *roc – auc* подсчитанной по макроусреднению в 98.61%. Что показывает существенный прирост точности с 97.05% для метода, использующего только видеофиксацию. Более того затраты времени на вычисления упали в десять раз и была исключена обработка данных на *GPU*, из за чего выросла скорость обработки с 16Гц до 100Гц.

4.2 Сравнительные характеристики для разных источников данных

По итогам работы сравнения результатов с моделями построенными на основе видео данных: были получены модели которые распознают по данным видео два действия человека, а именно, первое — поднятие руки выше 90° относительно туловища человека и покраска стены человеком. Алгоритмы обучались на двух наборах признаков, один из которых был масштабированным для потенциального изменения относительного размера человека на изображении. Для моделей распознавания рассматриваемых действий получены метрики ROC-AUC и Ассигасу на контрольных выборках со значениями порядка 91–99% и 80–98% соответственно.

В итоге сравнения алгоритмов производится на разных методах, но на данных, полученных в одном эксперименте как и в работе [116]. Так как обработка видео последовательности основывается на нейросетевом методе, а получение информации с помощью *MEMS* датчиков позволяет использовать алгоритмический метод получения результата, то второй подход дает более точные ре-

зультаты при значительно меньшей вычислительной сложности. Автором был разработан классификатор основанный на нейронных сетях на основе данных с датчиков. Результата основанного на методе *XGBoost* он не улучшил, однако вычислительных ресурсов требуется больше, и из-за недостаточности уделенного ему времени он был исключен из сравнения. Однако нейросетевой классификатор все так же остается возможным вариантом решения.

Для сравнения моделей были выбраны метрики ROC-AUC, и Accuracy. Основной метрикой бинарной классификации является ROC-AUC, так как она не искажается балансом классов. Однако в рамках рассматриваемой задачи также можно полагаться на метрику Accuracy, так как сильного дисбаланса классов на обучающей и контрольной выборках нет.

В общей сложности было получено 10 моделей 8 – по видео представлению, 2 – по IMU представлению.

Таблица 10 – Метрики *ROC – AUC* и *Accuracy* на контрольных данных для первой целевой переменной. (Поднятие руки)

Описание	<i>ROC – AUC</i> , %	<i>Accuracy</i> , %
Ракурс 1, Набор признаков 1	98.62	94.36
Ракурс 1, Набор признаков 2	98.56	94.90
Ракурс 2, Набор признаков 1	99.26	96.16
Ракурс 2, Набор признаков 2	99.20	95.91
3D модель по видео данным	98.91	94.78
IMU – данные	99.80	93.31

Рассматривая представленные результаты в таблицах 10, 11 можно увидеть, что для целевой переменной 1 ROC-AUC и Accuracy принципиально не различаются в зависимости от набора признаков, однако результаты получились лучше с IMU – данных, даже в сравнении с восстановленной 3D моделью. Полученные результаты обучения на первом целевом действии дают точность ROC AUC – 98% и Accuracy – 94% – по видео данным и ROC AUC – 96% и Accuracy – 93% на MEMS данных, на втором целевом действии – ROC AUC – 95% и Accuracy – 84% на видео данных и ROC AUC – 98% и Accuracy – 92% – на MEMS данных.

Таблица 11 — Метрики $ROC - AUC$ и $Accuracy$ на контрольных данных для второй целевой переменной (покраска стены).

Описание	$ROC - AUC, \%$	$Accuracy, \%$
Ракурс 1, Набор признаков 1	95.48	84.19
Ракурс 1, Набор признаков 2	94.25	82.42
Ракурс 2, Набор признаков 1	93.49	84.86
Ракурс 2, Набор признаков 2	91.75	80.54
3D модель по видео данным	94.69	86.35
IMU – данные	98.33	92.12

4.2.1 Результаты сравнения моделей для разных источников данных

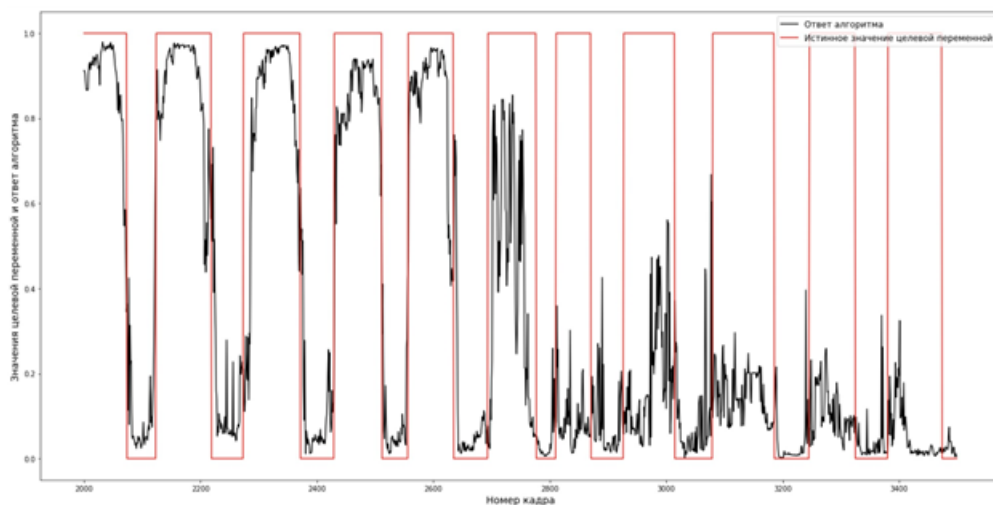


Рисунок 4.2 — Определение моделью целевой переменной 2 на данных с ракурса 2 и набором признаков 2 по видео.

На всех изображениях 4.2,4.4 можно увидеть, что предсказанные алгоритмом значения целевых переменных не на всех кадрах находятся близко к точным значениям целевых переменных. Например, такое явно можно заметить на рис. 4.2,4.4 на кадрах 2700–3500. Данные расхождения получаются из-за того, что на кадрах с видео рука человека перестает быть в кадре (заслоняется ту-

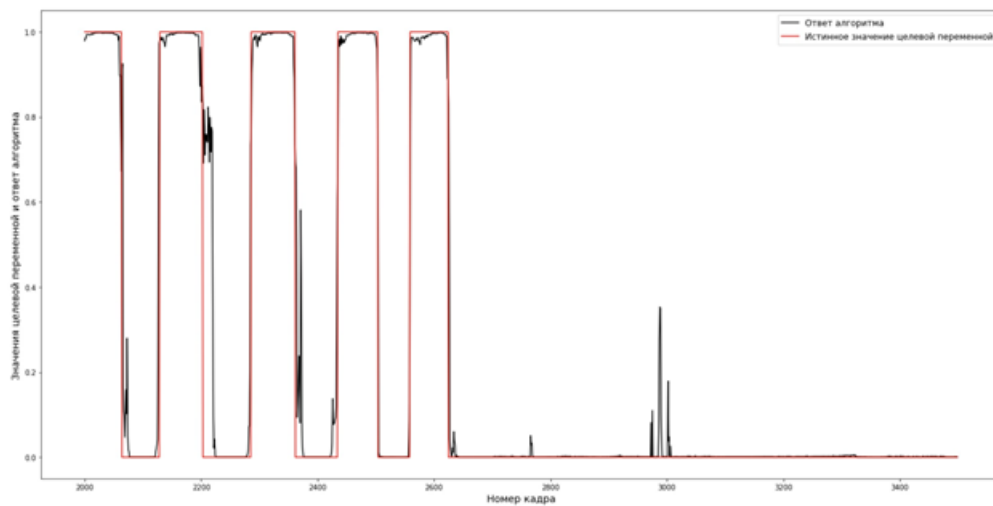


Рисунок 4.3 — Определение моделью целевой переменной 1 по MEMS данным.

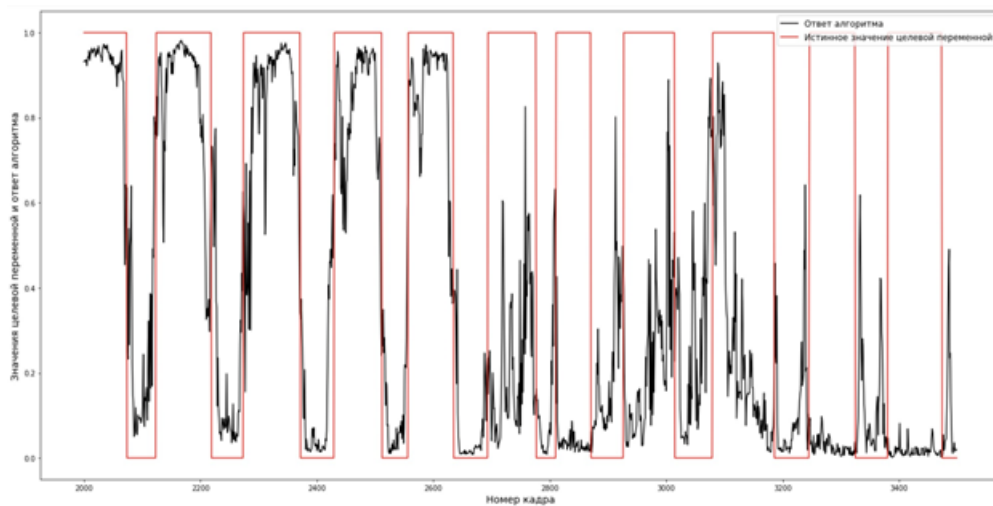


Рисунок 4.4 — Определение моделью целевой переменной 2 на данных с ракурса 2 и набором признаков 1 по видео данным.

ловищем человека, пример показан на рис. 4.5. Полученную проблему можно исправить, используя одновременно данные с двух камер.

Однако также на всех изображениях (4.2,4.4) видно, что на большей части кадров расхождения между истинным значением целевой переменной и вероятностью, полученной алгоритмом, достаточно малы, что как раз значит, что полученные алгоритмы в большинстве случаев справляются с определением истинных значений целевых переменных.

На рис. 4.3 показано определение первой целевой переменной (подъем руки) моделью на MEMS – данных, выходы как видно из графика менее шумные и так же достаточно точно определяют целевую переменную. На кадрах 2950-3050 на рис. 4.3 видны артефакты, они связаны с наклоном и подъемом предмета с пола человеком. На данных кадрах рука человека относительно тела находи-

лась около 90°, но грудная клетка человека располагалась в горизонтальном положении поэтому, данное действие не было размечено в выборке.

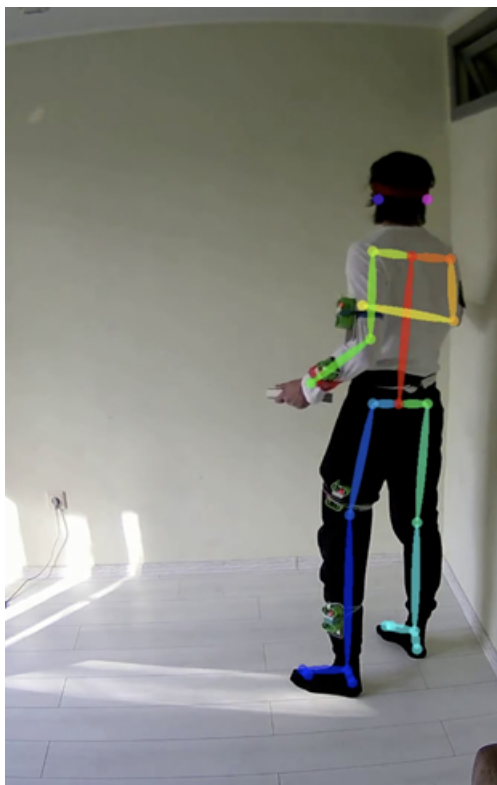


Рисунок 4.5 — Кадр, на котором туловище человека заслоняет руку

Для моделей, определяющих действие покраски стены, обученных на масштабированных признаках, рассматриваемые метрики получились ниже на 2–4% в сравнении с моделями, обученными на не масштабированных признаках. Это объясняется тем, что при масштабировании частично теряется информация о ключевых точках «скелета» человека при помощи системы OpenPose.

В заключении хотелось бы отметить, что при использовании MEMS-датчиков точность классификации последовательностей движений на основе данных эксперимента 3.3 приближается к 99%, что дает основание для создания промышленного инструмента на основе данной методики.

4.3 Прогноз вращений и ускорений для элементов модели тела человека

Для разработки устройств, применяемых в промышленных условиях и связанных с работой человека, необходимо иметь апробированную механическую модель системы, провести множество экспериментов в различных условиях. Все более распространенным становится использование современных технологий на промышленных производствах. В данной главе исследуется возможность предсказания кинематики движений человека в условиях промышленного производства на основе временных последовательностей. Целью исследования является разработка интеллектуальной системы управления экзоскелетом, которая анализирует механические действия человека и перераспределяет нагрузку на себя во время выполнения рабочих операций человеком. В предыдущих работах уже было предложено решение задачи предсказания момента начала следующего действия человека с высокой точностью. В данной задаче предлагается использовать модель основанную на CDPM для предсказания кинематических характеристик частей тела человека во время движения, используя информацию о предыдущих движениях, благодаря чему появится возможность сделать более подробные биомеханические исследования для анализ эргономики как рабочего места так и рабы в экзоскелете. Также этот подход может быть полезен для построения законов управления экзоскелетами например для работы [117] и повышения эффективности труда в промышленных условиях. Благодаря применению вариационных автоэнкодеров (VAE) [118] для лучшего воспроизведения временных последовательностей. Было показано что при достаточной обучающей выборке предложенный метод показывает хорошую точность по метрикам (CRPS) [119] – непрерывная ранжированная оценка вероятности и среднеквадратичная ошибка (MSE).

4.3.1 Постановка задачи

Задача состоит в том чтобы научиться предсказывать кинематические характеристики в ключевых точках тела человека, на основе данных о его движениях, в условиях промышленного производства с последующей целью создания интеллектуальной системы управления для экзоскелетов.

Существующие искусственные нейросети с хорошей точностью предсказывают 200мс временных последовательностей, чего явно недостаточно для построения надежной системы управления. В данном разделе описывается автором нейросетевая – модель для прогнозирования временных последовательностей, для частей тела человека более чем на 1200мс. Кватернионы 3.3.2 и вектора ускорений представляются в виде последовательности значений, где каждая компонента является отдельной размерностью временной последовательности. В данной задаче введено дополнительное ограничение, которое не только учитывает определенность следующего действия человека, но также использует информацию о предсказании времени начала следующего действия.

Так же учитывая проблему поднятую в работы [120] о мультимодальности такого рода архитектур было решено вместо большой нейронной сети для предсказания большего количества механических действий, что нейронные сети будут основаны на архитектуре D3VAE [121] с корректировками, для прогнозирования одной однотипной работы человека. В данной работе будет рассматриваться задача реконструкции повторяющихся шаблонов в неперiodических действиях.

Основным подходом служит метод описанный [120], с учетом предположений автора, специфики данных и доработками для применения на временных последовательностях ориентаций и ускорений полученных с IMU датчиков.

В данной задаче необходимо выполнить прогноз кинематических параметров, связанных с определенным движением одной из частей тела человека. Для этого используется информация о предполагаемом времени начала и завершения следующей последовательности движений, а также данные, полученные с MEMS датчиков, зафиксированные в предшествующий период движения.

Описанные известные характеристики полученные на основе данных с MEMS датчиков и уже содержат некоторую ошибку не только в методах, но

и из допущения, что датчик закреплен на части тела человека жестко, что в действительности практически невыполнимо. Точка крепления не неподвижна относительно костей и мышц человека, которые отвечают за кинематические характеристики частей тела, существенно влияющих на показания датчиков. Из специфики используемой модели следует, что подавать на вход необходимо предшествующую временную последовательность. Однако учитывая известную информацию о виде, действия времени его начала и времени его конца, из предположения что будет получена более точная модель, на вход в качестве априорных знаний будет подаваться не предшествующие кинематические характеристики по времени, а предшествующие кинематические характеристики предыдущего действия человека, что по предположению автора существенно улучшит работу модели. Стоит отметить что не ожидается попадать ровно в ту же фазу предшествующего действия, важно лишь то, чтобы действия совпадали. Другими словами можно было бы составить новое действие человека из паттернов предыдущего действия.

Стоит отметить, что все получаемые данные с IMU сенсоров шумят, однако очевидно, что все кинематические характеристики частей тела человека во время движения изменяются непрерывно. И именно для этих характеристик будет необходимо создавать ассистирующее силовой моментное управление со стороны экзоскелета. Таким образом для целевой функции Y в обучающей выборке можно брать сильно усредненные и аппроксимированные кинематические характеристики, чтобы получать менее шумящие результаты.

4.3.2 Описание эксперимента

Проведен эксперимент, в котором производилось наблюдение на 16-тью здоровых молодых мужчинах-добровольцах исследовалось влияние пассивного экзоскелета верхних конечностей на моторный контроль. Анализировались данные и параметры выполнения инструкции в задаче на точность при движении правой рукой с объектом (шуруповертом), имитирующим рабочий процесс. Оценивалась точность попадания по мишеням на моноблоке подвешенным в фиксированном положении над уровнем глаз таким образом, что каждый доброволец

был вынужден поднять голову под углом в 45 градусов к горизонту что бы смотреть точно в центр экрана, в свою очередь экран был повернут к вертикальной плоскости под 45 градусов. Каждого добровольца снимали две перпендикулярные камеры. Каждый доброволец был одет в костюм из IMU и EMG датчиков, с каждого из датчиков собиралась информации о кватернионе поворота локальной системы координат, линейном ускорении с вычтенным векторов гравитации и линейной скорости. Именно эти данные будут рассматриваться как набор обучающих примеров для предсказания движений в настоящей публикации.

4.3.3 Реализация описанного метода

Предлагается моделировать предсказание значения временной последовательности в каждый момент времени на основе временной последовательности предыдущего движения человека и текущего значения, предполагая, что текущие значения распределены вокруг среднего значения и могут быть представлены с помощью случайной величины, зависящей от предыдущего значения и временного интервала от начала механического действия человека. Таким образом, для каждого момента времени мы можем смоделировать процесс диффузии и получить новое значение временной последовательности на основе текущего значения и временного интервала прошлого механического действия. При этом новое значение может быть отлично относительно текущего значения, что позволяет генерировать разнообразные значения временной последовательности.

Сначала входные и выходные временные последовательности одновременно дополняются связанным процессом зашумления. Затем диффузированные входные временные последовательности подаются на вход модели BVAE – ограниченный вариационный автоэнкодер для получения целевого не шумного временной последовательности, который можно рассматривать как обратный процесс к зашумлению. Применяется механизм оценки качества исключения шума для того, чтобы полученный целевая последовательность двигался в сторону истинного целевой последовательности. В то же время скрытые состояния в BVAE используются для декомпозиции, что позволяет улучшить интерпретируемость и надежность модели.

Для временных последовательностей используется рекуррентный двунаправленный вариационный автоэнкодер (*BVAE*). *BVAE* имеет две основные части: кодировщик (*encoder*) и декодировщик (*decoder*). Кодировщик принимает на вход временная последовательность и преобразует его в латентное пространство, где каждый элемент латентного пространства представляет собой скрытый признак временной последовательности. Декодировщик затем преобразует эти скрытые признаки обратно во временная последовательность. Обе части модели оптимизируются одновременно, чтобы увеличить правдоподобие входного временной последовательности. *BVAE* использует механизм оценки качества исключения шума (*denoising score-matching*), который позволяет оценивать правдоподобие целевого распределения путем декорреляции шума во входных данных. Этот механизм улучшает качество генерации, поскольку он позволяет модели сконцентрироваться на основных характеристиках данных, не привязываясь к шуму. Тем самым создавая более точные признаки временной последовательности.

Таким образом входным вектором X будет 10-мерный вектор составленный из покомпонентной склейки одного четырехмерного временной последовательности кватернионов обезразмеренный и двух трехмерных временных последовательностей вектора направления гравитации и ускорений для одного датчика. Каждая компонента вектора обезразмерена на среднее значение на скользящем окне в сто последних значений. Целевой временной последовательности соответствует вектор Y . Предполагается, что Y может быть сгенерирован из скрытых переменных состояний $Z \in \Omega_Z$, которые могут быть получены из гауссовского распределения $Z \sim P(Z|X)$. Распределение скрытых переменных состояний может быть сформулировано как $P(Z|X) = g(X)$, где g обозначает нелинейную функцию. Тогда, плотность данной целевой последовательности задается следующим образом:

$$P(Y) = \int_{\Omega} P(Z|X)(Y - f(Y))dZ,$$

где f обозначает параметризованную функцию. Целевой временная последовательность может быть получен напрямую путем выборки из $P(Y)$.

В данной задаче прогнозирования временных последовательностей требуется научиться извлекать представление Z , которое улавливает полезные сиг-

налы из X , и отобразить многомерный вектор X в скрытое пространство ярко выражающее необходимое действие.

Для данного временной последовательности X и его внутреннего шума X_e , мы имеем разложение: $X = (X_r \circ X_e)$, где X_r - это идеальные данные временной последовательности без шума. X_r и X_e независимы друг от друга. Пусть $P(Z|X) = P(Z|X_r \circ X_e)$, тогда оценочная целевая последовательность Y_b может быть создан с распределением $P(Y|Z) = P(Y_{br}|Z)P(Y_{be}|Z)$, где Y_{br} - это идеальная часть Y_b , а Y_{be} - это оценочный шум. Без потери общности, модель может полностью улавливать Y_{br} . То есть, $\|Y_r - Y_{br}\|$, где Y_r - это идеальная часть целевой последовательности Y . Кроме того, Y может быть разложен как $Y = (Y_{br} \circ Y_e)$, Y_e обозначает шум Y . Таким образом, ошибка может быть оценена как сочетание алеаторической неопределенности и эпистемической неопределенности. Алеаторическая неопределенность измеряет то, что мы не можем понять из данных. Она могла бы быть исключена если бы была возможностью наблюдать все объясняющие переменные с повышенной точностью. Эпистемическая неопределенность измеряет то, чего наша модель не знает из-за недостатка обучающих данных.

В сравнение с предложенным подходом будут рассматриваться классический статистический подход на основе модели ARIMA и другой подход основанный на нейронных сетях авторегрессивная рекуррентная модель RNN. ARIMA (Autoregressive Integrated Moving Average) - это статистическая модель для анализа и прогнозирования временных последовательностей. Она объединяет авторегрессию (AR), интеграцию (I) и скользящее среднее (MA). AR - компонента отражает зависимость текущего значения последовательности от предыдущих значений. I - компонента используется для преобразования последовательности в стационарную, устраняя тренды и сезонность. MA-компонента учитывает случайные шоки, которые влияют на текущее значение последовательности. ARIMA модель имеет параметры, определяющие количество AR и MA компонентов, а также степень интеграции для достижения стационарности. ARIMA используется для исследования зависимостей в данных, а также для предсказания будущих значений последовательности на основе его прошлого поведения. Основная цель ARIMA - понять и описать временные последовательности с помощью статистических свойств и использовать их для прогнозирования.

Архитектура авторегрессивной RNN (рекуррентной нейронной сети) расширяет традиционную модель RNN, включая авторегрессивные компоненты, позволяя модели использовать свои собственные предыдущие прогнозы в качестве входных данных. Это позволяет модели учитывать как историческую информацию, так и обратную связь, создаваемую ее собственными предсказаниями. Такая модель улавливает временные связи и зависимости между различными компонентами временной последовательности. Подход состоит в обучении модели историческими данными движений для выявления закономерностей, и создания точных прогнозов для будущих временных шагов.

Для обработки одномерных временных последовательностей, содержащих кватернионы и ускорения, использовались сверточные слои с одномерными фильтрами. Эти фильтры выделяют определенные временные паттерны и зависимости в данных.

Значение длины прогноза, были выбраны согласно представлению о шаблонном действии человека, от 25 значений до 250 значений в зависимости от действия. На данный момент были проведены лабораторные испытания данной системы, обучены модели предсказания не периодических механических действий человека в условиях промышленного производства.

4.3.4 Результаты работы алгоритмов прогнозирования

В результате удается сгенерировать предсказание для многомерного временной последовательности состоящего из компонент вращения, линейной скорости и линейных ускорений. С точностями подсчитанными по метрикам средней квадратичной ошибки - 0.112 ± 0.033 , измеряемой как среднее значение квадрата разности между прогнозируемыми значениями модели и фактическими значениями целевой переменной. Однако эта метрика имеет квадратичную единицу измерения, не интерпретируемую в контексте задачи и по метрике CRPS - 0.228 ± 0.026 , которая измеряет интегральное расстояние между функцией прогнозирования вероятностей и функцией фактических вероятностей. Она учитывает не только среднее значение прогнозируемых вероятностей, но также и форму кривой вероятностей. Существует еще одна метрика для сравне-

ния временных последовательностей MRE она вычисляется путем сравнения прогнозируемого значения с соответствующим фактическим значением и определения относительной разницы между ними. Затем вычисляется средняя относительная разница для всех прогнозируемых значений. Эта метрика имеет физический смысл в том, что прогнозируемые значения отличаются от фактических значений в среднем на долю метрики отличной от нуля относительно единицы.

Таблица 12 – результаты сравнения статистических методов и подхода основанного на искусственных нейронных сетях. На основе сравнительной таблицы можно сделать вывод о том, что в целом статистические подходы основанные на максимальном правдоподобию показывают результат выше чем на основе методов оптимизации или метода наименьших квадратов, однако лидерами являются нейросетевые методы в рамках этой задачи.

Таблица 12 — Сравнительная таблица на основе метрик MSE и $CRPS$

Название метода	MSE	$CRPS$
ARIMA	1.015 ± 0.12	0.523 ± 0.1
SARIMA	0.838 ± 0.22	0.621 ± 0.11
NaiveForecaster	1.561 ± 0.04	0.9 ± 0.061
ExponentialSmoothing	2.121 ± 0.033	0.88 ± 0.26
RNN	0.771 ± 0.06	0.451 ± 0.055
D3VAE	0.112 ± 0.033	0.228 ± 0.026

Заключение

Основные результаты проведенного исследования:

Созданы способы определения динамических параметров и энергозатрат при движении сегментов тела человека с использованием микромеханических датчиков на основе биомеханической модели.

Дано теоретическое обоснование разработанным способам определения динамических характеристик.

Разработан метод полуавтоматической калибровки системы, состоящей из инерциальных датчиков и сегментов тела человека, на которых они размещены.

Разработана аппаратно-программная система, которая позволяет корректно собирать, хранить, калибровать, фильтровать данные с датчиков для последующей визуализации в модельной системе координат.

Выполнено сравнение алгоритмов определения последовательности движения на основе эксперимента с данными из разных источников с целью получения достоверных данных сравнение выполнено при одинаковых методиках получения конечного результата. На основе таблиц 10, 11 - можно сделать вывод что использование микромеханических датчиков дает существенный прирост в точности по метрике $roc - auc$ подсчитанной по макроусреднению в 98.61%. Более того затраты времени на вычисления упали в десять раз и была исключена обработка данных на *GPU*, из за чего выросла скорость обработки данных с 16Гц до 100Гц. Однако стоит отметить и ограничения на эргономику рабочего места и дополнительное время для начала трудового процесса, в которое необходимо включить калибровку датчиков. В среднем по результатам экспериментов это занимало у неподготовленных операторов порядка 20 ± 5 минут, у подготовленных 10 ± 3 .

Разработан алгоритм автоматического восстановления трехмерных данных о движении сегментов тела человека (скелетных данных) на основе видеоданных. Предложенный алгоритм позволяет избегать вырождения при любых положениях камеры и извлекать трехмерную информацию о движении человека из двух и более двумерных некомпланарных изображений.

Разработана система моделей для анализа последовательностей движений человека. В рамках данной системы были построены методы машинного обуче-

ния на основе данных микромеханических датчиков и данных видео фиксации. Методы способны идентифицировать текущую последовательность движений (действие) человека, предсказывать последующие действие и определять момент начала последующего действия. Такая система может быть использована в различных областях, включая разработку интерактивных систем виртуальной и дополненной реальности, контроль за движением в спорте или медицине.

Разработаны программные средства для прогнозирования положений и ускорений сегментов тела человека на 1 секунду вперед, которая учитывает последовательность движений сегментов тела человека в предыдущие моменты времени. Для предсказания используется нейросеть. Результаты экспериментов показывают достаточно хорошее совпадение предсказания с измерениями. Это позволяет сделать вывод о том, что нейросеть позволяет успешно выполнять экстраполяцию регулярных – не случайных, движений сложной механической системы.

Проведен лабораторный эксперимент по апробации системы моделей для анализа последовательностей движений человека. В экспериментах получена достаточно хорошая точность по метрике *accuracy* для классификации — 99%, предсказание следующего действия — 93% и предсказание момента начала следующего действия — 91%.

Список сокращений и условных обозначений

МО – машинное обучение, чаще употребляется как модель машинного обучения.

MEMS – (Micro-Electro-Mechanical Systems) это многопроцессорная и мультифункциональная интеграция микро и наномасштабных механических и электронных компонентов на одном полупроводниковом подложке, что позволяет создавать микроустройства, способные к детектированию, управлению и взаимодействию с физическим окружением на микроуровне. Эти системы могут включать в себя микроактуаторы, микросенсоры, микроэлектроды и другие компоненты, обеспечивая широкий спектр приложений, включая медицину, электронику, авиацию, и многие другие области.

CNN – (Convolutional Neural Network) это класс нейронных сетей, который широко используется в обработке изображений и анализе данных, где пространственная структура данных имеет важное значение. CNN специализируются на обнаружении и извлечении важных признаков из изображений и последовательностей данных, таких как видео и звук. Они получили широкое применение в областях компьютерного зрения, распознавания образов, анализа временных рядов и многих других задач.

RNN – (Recurrent Neural Network) Рекуррентная нейронная сеть это класс искусственных нейронных сетей, в которых нейроны обладают внутренней памятью и способностью передавать информацию из предыдущих шагов времени к текущему. Это позволяет RNN моделировать и анализировать последовательные данные, сохраняя контекст и учитывая зависимости между элементами последовательности. RNN применяются в различных задачах, включая временные ряды, обработку естественного языка, машинный перевод и другие приложения, где важно учитывать последовательность данных.

LSTM – (Long Short-Term Memory) это вид рекуррентной нейронной сети (RNN), который спроектирован для решения проблемы затухающих и взрывающихся градиентов при обучении на последовательных данных. LSTM представляет собой архитектуру, обладающую внутренними ячейками и механизмами входных и выходных ворот, которые позволяют эффективно управлять и передавать информацию через различные временные шаги. Этот тип нейронных

сетей широко используется для анализа временных рядов, обработки текста, генерации последовательностей и других задач, где важна учетная долгосрочной зависимости данных.

OVR – (One-Versus-Rest) это метод бинарной классификации, используемый в машинном обучении, при котором каждый класс сравнивается с остальными классами отдельно. Другими словами, для задачи с множеством классов создается отдельный классификатор для каждого класса, который обучается отличать этот класс от всех остальных. Когда необходимо предсказать класс для нового примера, каждый из классификаторов возвращает свой результат, и класс с наибольшим уверенностью (вероятностью) считается предсказанным классом.

MНК – Метод наименьших квадратов это статистический метод, который используется для нахождения параметров (коэффициентов) линейной модели или уравнения так, чтобы минимизировать сумму квадратов отклонений (остатков) между наблюдаемыми значениями зависимой переменной и значениями, предсказанными этой моделью. МНК широко используется в регрессионном анализе и линейной аппроксимации данных и может быть расширен на более сложные модели.

В контексте регрессии, МНК позволяет найти линию (или гиперплоскость в многомерном пространстве), которая наилучшим образом соответствует набору данных, минимизируя сумму квадратов вертикальных расстояний между точками данных и линией. Этот метод применяется во многих областях, включая экономику, инженерию, физику, биологию и многие другие, для анализа и моделирования зависимостей между переменными и прогнозирования результатов.

IMU – это электронное устройство, состоящее из датчиков, предназначенных для измерения и регистрации параметров движения и ориентации в трехмерном пространстве. Типичные компоненты IMU включают акселерометры, которые измеряют ускорение объекта, гироскопы, которые измеряют угловую скорость вращения, и магнитометры, которые помогают определить ориентацию в магнитном поле Земли.

NDOF – Режим работы NDOF (Nine Degrees of Freedom) в датчике BNO055 представляет собой один из режимов работы этого датчика. BNO055 – это инерциальная модульная единица (IMU), которая интегрирует в себе ак-

селерометр, гироскоп и магнитометр, позволяя измерять параметры движения и ориентации объекта в трехмерном пространстве. Режим NDOF – это наиболее продвинутый режим работы BNO055 и обеспечивает измерение всех девяти степеней свободы (nine degrees of freedom):

СК – система координат это математический и графический способ описания и определения положения точек и объектов в пространстве или на плоскости. С помощью системы координат можно задать числовые координаты (обычно две или три) для точек в пространстве, что позволяет однозначно определить их положение.

TP – True Positives это количество положительных примеров, которые модель правильно классифицировала как положительные. Это означает, что модель верно обнаружила их как членов целевого класса.

FP – False Positives это количество отрицательных примеров, которые модель неправильно классифицировала как положительные. Это означает, что модель ложно считала их членами целевого класса.

TN – True Negatives это количество отрицательных примеров, которые модель правильно классифицировала как отрицательные. Это означает, что модель верно определила их как не членов целевого класса.

FN – False Negatives это количество положительных примеров, которые модель неправильно классифицировала как отрицательные. Это означает, что модель упустила их как членов целевого класса.

EMG – Электромиография это метод исследования, который использует электроды для регистрации электрических сигналов, генерируемых мышцами при их сокращении и расслаблении. Эти сигналы записываются и анализируются с целью оценки мышечной активности, выявления патологий в мышцах и нервной системе, а также для биомеханических исследований движения. EMG широко используется в медицинских диагностических и терапевтических целях, а также в научных исследованиях и реабилитации пациентов.

MSE – это метрика, используемая для измерения средней квадратичной ошибки между фактическими и предсказанными значениями в регрессионных моделях. Она вычисляется как средний квадрат разницы между фактическими значениями (обычно обозначаемыми как y_i) и предсказанными значениями (обычно обозначаемыми как \hat{y}_i) для всех наблюдений i :

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

где n – количество наблюдений. Чем меньше значение MSE, тем лучше модель соответствует данным.

CRPS – это метрика, используемая для оценки качества прогнозов вероятностей в задачах прогнозирования вероятностных распределений. Она измеряет расстояние между эмпирическим распределением фактических наблюдений и предсказанным вероятностным распределением. CRPS вычисляется как интеграл от квадрата разности между функциями распределения фактических и предсказанных значений:

$$CRPS = \int_{-\infty}^{\infty} [F_{\text{observed}}(x) - F_{\text{predicted}}(x)]^2 dx$$

где $F_{\text{observed}}(x)$ – функция распределения фактических значений, $F_{\text{predicted}}(x)$ – функция распределения предсказанных значений. Меньшее значение CRPS указывает на более точные и надежные вероятностные прогнозы.

Словарь терминов

Kinect – это коммерческая линейка датчиков и сенсорных устройств, разработанная и выпущенная компанией Microsoft. Kinect предназначен для захвата движений человека и распознавания голосовых команд в интерактивных приложениях и играх. Оно представляет собой комплексный системный модуль

skeleton – это специализированный формат файла, используемый для хранения информации о трехмерных скелетных моделях, обычно в контексте компьютерной графики и анимации. Этот формат предназначен для представления анимированных человеческих или животных персонажей, а также для захвата и сохранения данных о движении и анимации.

Accuracy – это мера точности классификационной модели, которая показывает долю правильных предсказаний, сделанных моделью относительно общего числа предсказаний. Она вычисляется следующим образом:

$$\text{Accuracy} = \frac{\text{Количество правильных предсказаний}}{\text{Общее количество предсказаний}}$$

Accuracy выражается в виде десятичной дроби или процентов и предоставляет информацию о том, насколько хорошо модель справляется с задачей классификации. Чем выше значение Accuracy, тем более точной считается модель.

ROC AUC – (Receiver Operating Characteristic Area Under the Curve) это метрика, используемая для оценки качества бинарной классификационной модели. Она измеряет площадь под кривой ROC (Receiver Operating Characteristic), которая представляет собой графическую зависимость между долей истинно положительных предсказаний (True Positive Rate) и долей ложно положительных предсказаний (False Positive Rate) при различных порогах классификации.

Кривая ROC строится следующим образом. Для каждого возможного порога классификации модель вычисляет True Positive Rate (TPR) и False Positive Rate (FPR). TPR – это доля верно классифицированных положительных примеров среди всех истинно положительных примеров, а FPR – это доля неверно классифицированных отрицательных примеров среди всех истинно отрицатель-

ных примеров. Постепенно изменяя порог классификации, мы получаем набор точек на плоскости (FPR, TPR), которые образуют кривую ROC.

ROC AUC – это числовая метрика, которая показывает площадь, заключенную между этой кривой и нулевой линией (кривой случайного классификатора). Значение ROC AUC находится в диапазоне от 0 до 1, где 0.5 соответствует случайному классификатору, а 1.0 – идеальной классификации.

Формула для вычисления ROC AUC:

$$\text{ROC AUC} = \int_0^1 \text{True Positive Rate } d(\text{False Positive Rate})$$

Чем выше значение ROC AUC, тем лучше работает модель в задаче бинарной классификации. ROC AUC позволяет сравнивать и оценивать производительность различных классификационных моделей, учитывая компромисс между TPR и FPR при различных порогах.

ARIMA – (Autoregressive Integrated Moving Average) это статистический метод анализа временных рядов, который используется для прогнозирования будущих значений на основе исторических данных. Модель ARIMA объединяет в себе три основных компонента: AR (Autoregressive): Этот компонент представляет собой авторегрессию и описывает зависимость текущего значения временного ряда от его предыдущих значений. В ARIMA модели, это выражается через авторегрессионный параметр p , который указывает количество лагов (предыдущих моментов времени), используемых для прогнозирования текущего значения. I (Integrated): Этот компонент описывает степень дифференцирования временного ряда, необходимую для стабилизации его статистических свойств. Если временной ряд нестационарен (имеет тренд или сезонность), то проводится дифференцирование для превращения его в стационарный временной ряд. Параметр d указывает количество раз, которое необходимо выполнить дифференцирование. MA (Moving Average): Этот компонент представляет собой скользящее среднее и описывает зависимость текущего значения временного ряда от предыдущих ошибок прогнозирования. В ARIMA модели, это выражается через параметр q , который указывает количество лагов ошибок, используемых для прогнозирования текущего значения.

Список литературы

1. *Аршавский Ю.* ИМ Гельфанд о математике и нейрофизиологии // Вестник Российской академии наук. — 2010. — т. 80, № 10. — с. 937—941.
2. A new multi-position calibration method for MEMS inertial navigation systems / Z. F. Syed [и др.] // Measurement science and technology. — 2007. — т. 18, № 7. — с. 1897.
3. Topical antifungal treatments for tinea cruris and tinea corporis / M. El-Gohary [и др.] // Cochrane Database of Systematic Reviews. — 2014. — № 8.
4. *El-Gohary M., Pearson S., McNames J.* Joint angle tracking with inertial sensors // 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. — IEEE. 2008. — с. 1068—1071.
5. A novel calibration method of magnetic compass based on ellipsoid fitting / J. Fang [и др.] // IEEE Transactions on Instrumentation and Measurement. — 2011. — т. 60, № 6. — с. 2053—2061.
6. Adaptive reconstruction of human motion on wireless body sensor networks / S.-Y. Chen [и др.] // 2011 International Conference on Wireless Communications and Signal Processing (WCSP). — 2011. — с. 1—5. — DOI: [10.1109/WCSP.2011.6096954](https://doi.org/10.1109/WCSP.2011.6096954).
7. An enhanced multi-position calibration method for consumer-grade inertial measurement units applied and tested / T. Nieminen [и др.] // Measurement Science and Technology. — 2010. — т. 21, № 10. — с. 105204.
8. Online IMU self-calibration for visual-inertial systems / Y. Xiao [и др.] // Sensors. — 2019. — т. 19, № 7. — с. 1624.
9. Sensor-to-segment calibration methodologies for lower-body kinematic analysis with inertial sensors: A systematic review / L. Pacher [и др.] // Sensors. — 2020. — т. 20, № 11. — с. 3322.

10. Body Pose Prediction Based on Motion Sensor Data and Recurrent Neural Network / M. Woźniak [и др.] // IEEE Transactions on Industrial Informatics. — 2021. — т. 17, № 3. — с. 2101–2111. — DOI: [10.1109/TII.2020.3015934](https://doi.org/10.1109/TII.2020.3015934).
11. Predicting human movement with multiple accelerometers using movelets / B. He [и др.] // Medicine and science in sports and exercise. — 2014. — т. 46, № 9. — с. 1859.
12. *Ashby B. M., Delp S. L.* Optimal control simulations reveal mechanisms by which arm movement improves standing long jump performance // Journal of biomechanics. — 2006. — т. 39, № 9. — с. 1726–1734.
13. *Blemker S. S., Delp S. L.* Three-dimensional representation of complex muscle architectures and geometries // Annals of biomedical engineering. — 2005. — т. 33. — с. 661–673.
14. An interactive graphics-based model of the lower extremity to study orthopaedic surgical procedures / S. L. Delp [и др.] // IEEE Transactions on Biomedical engineering. — 1990. — т. 37, № 8. — с. 757–767.
15. *Horsman M. K.* The Twente lower extremity model. Consistent dynamic simulation of the human locomotor apparatus. — 2007.
16. A model of the lower limb for analysis of human movement / E. M. Arnold [и др.] // Annals of biomedical engineering. — 2010. — т. 38. — с. 269–279.
17. TLEM 2.0—A comprehensive musculoskeletal geometry dataset for subject-specific modeling of lower extremity / V. Carbone [и др.] // Journal of biomechanics. — 2015. — т. 48, № 5. — с. 734–741.
18. OpenSim: Open-source Software to Create and Analyze Dynamic Simulations of Movement / S. L. Delp [и др.] // IEEE Transactions on Biomedical Engineering. — 2007. — т. 54, № 11. — с. 1940–1950. — DOI: [10.1109/TBME.2007.901024](https://doi.org/10.1109/TBME.2007.901024).
19. Analysis of musculoskeletal systems in the AnyBody Modeling System / M. Damsgaard [и др.] // Simulation Modelling Practice and Theory. — 2006. — т. 14, № 8. — с. 1100–1111.
20. Creating interactive virtual humans: Some assembly required / J. Gratch [и др.] // IEEE Intelligent systems. — 2002. — т. 17, № 4. — с. 54–63.

21. *De Sevin E., Kallmann M., Thalmann D.* Towards real time virtual human life simulations // Proceedings. Computer Graphics International 2001. — IEEE. 2001. — с. 31—37.
22. *Кий К., Анохин Д., Подопросветов А.* Программная система обработки изображений с параллельными вычислениями // Программирование. — 2020. — № 6. — с. 41—54.
23. *Komura T., Shinagawa Y., Kunii T. L.* Creating and retargetting motion by the musculoskeletal human body model // The visual computer. — 2000. — т. 16. — с. 254—270.
24. *Delp S. L., Loan J. P.* A graphics-based software system to develop and analyze models of musculoskeletal structures // Computers in biology and medicine. — 1995. — т. 25, № 1. — с. 21—34.
25. *Delp S. L., Loan J. P.* A computational framework for simulating and analyzing human and animal movement // Computing in Science & Engineering. — 2000. — т. 2, № 5. — с. 46—55.
26. *Suzuki N., Takatsu A.* 3d and 4d visualization of Morphological and Functional Information from the Human Body using Noninvasive Measurement Data citation. Atlas of visualization. — 1997.
27. *Nakamura Y., Yamane K.* Dynamics computation of structure-varying kinematic chains and its application to human figures // IEEE Transactions on Robotics and Automation. — 2000. — т. 16, № 2. — с. 124—134.
28. Upper limb kinematics using inertial and magnetic sensors: Comparison of sensor-to-segment calibrations / B. Bouvier [и др.] // Sensors. — 2015. — т. 15, № 8. — с. 18813—18833.
29. Realtime multi-person 2d pose estimation using part affinity fields / Z. Cao [и др.] // Proceedings of the IEEE conference on computer vision and pattern recognition. — 2017. — с. 7291—7299.
30. Towards 3d human pose estimation in the wild: a weakly-supervised approach / X. Zhou [и др.] // Proceedings of the IEEE international conference on computer vision. — 2017. — с. 398—407.

31. Xsens MVN: Full 6DOF human motion tracking using miniature inertial sensors / D. Roetenberg, H. Luinge, P. Slycke [и др.] // Xsens Motion Technologies BV, Tech. Rep. — 2009. — т. 1. — с. 1—7.
32. Официальный сайт компании Xsens Technologies В. В. Электронный ресурс. // — URL: <https://www.xsens.com/motion-capture>.
33. *Henk J. L. R. J. S.* Motion tracking system: patent no. US8165844B2 US. — 2007.
34. Официальный сайт компании ООО "Неврокор" Электронный ресурс. // — URL: <http://www.neurocor.ru/>.
35. *Park H. S., Shi J.* Force from Motion: Decoding Control Force of Activity in a First-Person Video // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2018. — т. 42, № 3. — с. 622—635.
36. Human body measurement estimation with adversarial augmentation / N. Ruiz [и др.] // 2022 International Conference on 3D Vision (3DV). — IEEE. 2022. — с. 219—230.
37. Learning realistic human reposing using cyclic self-supervision with 3d shape, pose, and appearance consistency / S. Sanyal [и др.] // Proceedings of the IEEE/CVF International Conference on Computer Vision. — 2021. — с. 11138—11147.
38. *Самойленко М.* Восстановление формы трехмерного объекта по двумерным изображениям // Научно-технический вестник информационных технологий, механики и оптики. — 2017. — т. 17, № 6. — с. 1074—1083.
39. *Тупицын И. В.* Реконструкция трехмерной модели объекта на основе стереопары при решении задач 3D-моделирования // Сибирский аэрокосмический журнал. — 2011. — 3 (36). — с. 88—92.
40. *Lara O. D., Labrador M. A.* A survey on human activity recognition using wearable sensors // IEEE communications surveys & tutorials. — 2012. — т. 15, № 3. — с. 1192—1209.
41. Deep learning for sensor-based activity recognition: A survey / J. Wang [и др.] // Pattern recognition letters. — 2019. — т. 119. — с. 3—11.

42. Adaptive sliding window segmentation for physical activity recognition using a single tri-axial accelerometer / M. H. M. Noor [и др.] // *Pervasive and Mobile Computing*. — 2017. — т. 38. — с. 41–59.
43. *Zheng X., Wang M., Ordieres-Meré J.* Comparison of data preprocessing approaches for applying deep learning to human activity recognition in the context of industry 4.0 // *Sensors*. — 2018. — т. 18, № 7. — с. 2146.
44. *Khan N. S., Ghani M. S.* A survey of deep learning based models for human activity recognition // *Wireless Personal Communications*. — 2021. — т. 120, № 2. — с. 1593–1635.
45. Temporal activity detection in untrimmed videos with recurrent neural networks / A. Montes [и др.] // *arXiv preprint arXiv:1608.08128*. — 2016.
46. Detecting the starting frame of actions in video / I. Kwak [и др.] // *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. — 2020. — с. 489–497.
47. Online human action detection using joint classification-regression recurrent neural networks / Y. Li [и др.] // *Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part VII 14*. — Springer. 2016. — с. 203–220.
48. *Wang B., Hoai M.* Back to the beginning: Starting point detection for early recognition of ongoing human actions // *Computer Vision and Image Understanding*. — 2018. — т. 175. — с. 24–31.
49. OpenPose [электронный ресурс]. — URL: <https://github.com/CMU-Perceptual-Computing-Lab/openpose>.
50. Instructing people for training gestural interactive systems / S. Fothergill [и др.] // *Proceedings of the SIGCHI conference on human factors in computing systems*. — 2012. — с. 1737–1746.
51. *Chen C., Jafari R., Kehtarnavaz N.* UTD-MHAD: A multimodal dataset for human action recognition utilizing a depth camera and a wearable inertial sensor // *2015 IEEE International conference on image processing (ICIP)*. — IEEE. 2015. — с. 168–172.
52. RGB-D-based action recognition datasets: A survey / J. Zhang [и др.] // *Pattern Recognition*. — 2016. — т. 60. — с. 86–105.

53. Ntu rgb+ d 120: A large-scale benchmark for 3d human activity understanding / J. Liu [и др.] // IEEE transactions on pattern analysis and machine intelligence. — 2019. — т. 42, № 10. — с. 2684—2701.
54. Human Activity Detection from RGBD Images. / J. Sung [и др.] // plan, activity, and intent recognition. — 2011. — т. 64.
55. *Koppula H. S., Gupta R., Saxena A.* Learning human activities and object affordances from rgb-d videos // The International journal of robotics research. — 2013. — т. 32, № 8. — с. 951—970.
56. *Xia L., Chen C.-C., Aggarwal J. K.* View invariant human action recognition using histograms of 3d joints // 2012 IEEE computer society conference on computer vision and pattern recognition workshops. — IEEE. 2012. — с. 20—27.
57. Towards understanding action recognition / H. Jhuang [и др.] // Proceedings of the IEEE international conference on computer vision. — 2013. — с. 3192—3199.
58. HMDB: a large video database for human motion recognition / H. Kuehne [и др.] // 2011 International conference on computer vision. — IEEE. 2011. — с. 2556—2563.
59. Lower Limb Exoskeleton with Hybrid Pneumatically Assisted Electric Drive: Models / I. Orlov [и др.] // International Conference on Robotics in Alpe-Adria Danube Region. — Springer. 2017. — с. 746—755.
60. Экзоскелет нижних конечностей с гибридным электро-пневматическим приводом. Модели и управление / И. Орлов [и др.] // Робототехника и техническая кибернетика. — 2015. — № 3. — с. 16—21.
61. *Орлов И. А., Грибков Д. А., Подопросветов А. В.* Промышленный экзоскелет нижних конечностей: конструкция и модели // Перспективные направления развития отечественных информационных технологий. — 2016. — с. 79—80.
62. Динамическая модель промышленного экзоскелета нижней части тела / И. Орлов [и др.] // Конференция проводится при финансовой поддержке Российского фонда фундаментальных исследований (проект № 18-01-20058). — с. 311.

63. Разработка системы управления биомеханическим комплексом помощи работникам производственных предприятий / А. Подопрсветов [и др.] // XII Всероссийский съезд по фундаментальным проблемам теоретической и прикладной механики. — 2019. — с. 196—198.
64. Прочностный расчет экзоскелета нижних конечностей человека / И. Орлов [и др.] // XXXI Международная инновационная конференция молодых ученых и студентов по проблемам машиноведения (МИК-МУС-2019). — 2020. — с. 156—159.
65. Оптимизация массы экзоскелета нижних конечностей человека / И. Орлов [и др.] // XXXI Международная инновационная конференция молодых ученых и студентов по проблемам машиноведения (МИК-МУС-2019). — 2020. — с. 152—155.
66. IMU-based estimation of lower limb motion trajectory with graph convolution network / Y.-L. Chen [и др.] // IEEE Sensors Journal. — 2021. — т. 21, № 21. — с. 24549—24557.
67. *Martinez J., Black M. J., Romero J.* On human motion prediction using recurrent neural networks // Proceedings of the IEEE conference on computer vision and pattern recognition. — 2017. — с. 2891—2900.
68. *Lehrmann A. M., Gehler P. V., Nowozin S.* Efficient nonlinear markov models for human motion // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. — 2014. — с. 1314—1321.
69. *Taylor G. W., Hinton G. E.* Factored conditional restricted Boltzmann machines for modeling motion style // Proceedings of the 26th annual international conference on machine learning. — 2009. — с. 1025—1032.
70. Recurrent network models for human dynamics / К. Fragkiadaki [и др.] // Proceedings of the IEEE international conference on computer vision. — 2015. — с. 4346—4354.
71. Structural-rnn: Deep learning on spatio-temporal graphs / А. Jain [и др.] // Proceedings of the IEEE conference on computer vision and pattern recognition. — 2016. — с. 5308—5317.

72. The pose knows: Video forecasting by generating pose futures / J. Walker [и др.] // Proceedings of the IEEE international conference on computer vision. — 2017. — с. 3332—3341.
73. Learning progressive joint propagation for human motion prediction / Y. Cai [и др.] // Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part VII 16. — Springer. 2020. — с. 226—242.
74. Skeleton-parted graph scattering networks for 3d human motion prediction / M. Li [и др.] // European Conference on Computer Vision. — Springer. 2022. — с. 18—36.
75. Msr-gcn: Multi-scale residual graph convolution networks for human motion prediction / L. Dang [и др.] // Proceedings of the IEEE/CVF International Conference on Computer Vision. — 2021. — с. 11467—11476.
76. Human3. 6m: Large scale datasets and predictive methods for 3d human sensing in natural environments / C. Ionescu [и др.] // IEEE transactions on pattern analysis and machine intelligence. — 2013. — т. 36, № 7. — с. 1325—1339.
77. Transformer networks for trajectory forecasting / F. Giuliari [и др.] // 2020 25th international conference on pattern recognition (ICPR). — IEEE. 2021. — с. 10335—10342.
78. *Ndahimana D., Kim E.-K.* Measurement methods for physical activity and energy expenditure: a review // Clinical nutrition research. — 2017. — т. 6, № 2. — с. 68—80.
79. OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields / Z. Cao [и др.] // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2019. — с. 75—77.
80. Currently OpenPose: Real-time multi-person keypoint detection library for body, face, hands, and foot estimation / H. Gines [и др.] //. — URL: <https://github.com/CMU-Perceptual-Computing-Lab/openpose>.

81. Development of complex control system for the autonomous vehicle Niva / A. Podoprosvetov [и др.] // 2019 XXI International Conference Complex Systems: Control and Modeling Problems (CSCMP). — IEEE. 2019. — с. 311—315.
82. *Huang J., Ling C. X.* Using AUC and accuracy in evaluating learning algorithms // IEEE Transactions on Knowledge and Data Engineering. — 2005. — т. 17, № 3. — с. 299—310.
83. *Google.* Receiver Operating Characteristic (ROC) and Area Under the Curve (AUC). — 2021. — URL: <https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc?hl=ru> ; Accessed: 2023.
84. *Olah C.* Understanding LSTM Networks. — 2015. — URL: <https://colah.github.io/posts/2015-08-Understanding-LSTMs/> ; Accessed: Month Day, Year.
85. *Chua L. O., Roska T.* The CNN paradigm // IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications. — 1993. — т. 40, № 3. — с. 147—156.
86. Применение нейронных сетей для предсказания физических действий человека / А. В. Подопроеветов [и др.] // Препринты Института прикладной математики им. МВ Келдыша РАН. — 2021. — № 109. — с. 1—16.
87. Sensor-to-Segment Calibration Methodologies for Lower-Body Kinematic Analysis with Inertial Sensors: A Systematic Review / L. Pacher [et al.] // Sensors. — 2020. — Vol. 20. — P. 3322. — DOI: [10.3390/s20113322](https://doi.org/10.3390/s20113322)..
88. Full body musculoskeletal model for muscle-driven simulation of human gait / A. Rajagopal [и др.] // IEEE Transactions on Biomedical Engineering. — 2016. — с. 1—1. — DOI: [10.1109/TBME.2016.2586891](https://doi.org/10.1109/TBME.2016.2586891).
89. Алгоритм калибровки группы MIMU датчиков / А. В. Подопроеветов [и др.] // Препринты Института прикладной математики им. МВ Келдыша РАН. — 2021. — № 43. — с. 1—16.
90. *Raviteja V., Felipe A., Rama C.* R3DG features: Relative 3D geometry-based skeletal representations for human action recognition // Computer Vision and Image Understanding. — 2016. — т. 152. — с. 155—166.

91. *Koelewijn A. D., Heinrich D., Bogert A. J. van den.* Metabolic cost calculations of gait using musculoskeletal energy models, a comparison study // *PloS one.* — 2019. — т. 14, № 9. — e0222037.
92. *Miller R. H.* A comparison of muscle energy models for simulating human walking in three dimensions // *Journal of biomechanics.* — 2014. — т. 47, № 6. — с. 1373—1381.
93. Искусственные мышцы (обзорная статья) / А. Алисейчик [и др.] // *Известия российской академии наук. теория и системы управления.* — № 2. — с. 137—161.
94. *Romero F., Alonso F.* A comparison among different Hill-type contraction dynamics formulations for muscle force estimation // *Mechanical Sciences.* — 2016. — т. 7, № 1. — с. 19—29.
95. *Bröde P., Kampmann B.* Accuracy of metabolic rate estimates from heart rate under heat stress—an empirical validation study concerning ISO 8996 // *Industrial health.* — 2019. — т. 57, № 5. — с. 615—620.
96. Flexing computational muscle: modeling and simulation of musculotendon dynamics / М. Millard [и др.] // *Journal of biomechanical engineering.* — 2013. — т. 135, № 2. — с. 021005.
97. *Umberger B. R.* Stance and swing phase costs in human walking // *Journal of the Royal Society Interface.* — 2010. — т. 7, № 50. — с. 1329—1340.
98. Removing the thermal component from heart rate provides an accurate $\dot{V}O_2$ estimation in forest work / P.-A. Dubé [и др.] // *Applied ergonomics.* — 2016. — т. 54. — с. 148—157.
99. Актуальность использования промышленных экзоскелетов для снижения количества профессиональных заболеваний опорно-двигательного аппарата верхней части тела / И. А. Орлов [и др.] // *Медицина труда и промышленная экология.* — 2019. — т. 59, № 7. — с. 412—416.
100. *Bhargava L. J., Pandy M. G., Anderson F. C.* A phenomenological model for estimating metabolic energy consumption in muscle contraction // *Journal of biomechanics.* — 2004. — т. 37, № 1. — с. 81—88.

101. *Брель Ю.* Взаимодействие и адаптация систем энергообеспечения скелетных мышц при физических нагрузках // Проблемы здоровья и экологии. — 2014. — 3 (41). — с. 47—53.
102. GitHub — веб-сервис для хостинга IT-проектов и их совместной разработки. Электронный ресурс. //. — URL: https://github.com/llecxis/Example%5C_IMUCalib.
103. *Caruso L., Russo R., Savino S.* Microsoft Kinect V2 vision system in a manufacturing application // Robotics and Computer-Integrated Manufacturing. — 2017. — т. 48. — с. 174—181. — DOI: <https://doi.org/10.1016/j.rcim.2017.04.001>.
104. *Chen T., Guestrin C.* XGBoost: A Scalable Tree Boosting System // CoRR. — 2016. — т. abs/1603.02754. — arXiv: [1603.02754](https://arxiv.org/abs/1603.02754). — URL: <http://arxiv.org/abs/1603.02754>.
105. *Chen C.* UTD-MHAD: A Multimodal Dataset for Human Action Recognition Utilizing a Depth Camera and a Wearable Inertial Sensor //. — 09.2015. — DOI: [10.1109/ICIP.2015.7350781](https://doi.org/10.1109/ICIP.2015.7350781).
106. NTU RGB+D 120: A large-scale benchmark for 3D human activity understanding / J. Liu [и др.] // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2020. — т. 42, № 10. — с. 2684—2701.
107. *Шматова Е. В.* Применение непрерывной логики для определения границ выбросов // Вестник КРАУНЦ. Физ.-мат. науки. — 2018. — № 3. — с. 190—194. — DOI: [10.18454/2079-6641-2018-23-3-190-194](https://doi.org/10.18454/2079-6641-2018-23-3-190-194).
108. Mahalanobis Distance. — URL: https://ru.frwiki.wiki/wiki/Distance_de_Mahalanobis ; (Accessed: 2022).
109. *Астафьева Н. М.* Вейвлет-анализ: основы теории и примеры применения // Успехи физических наук. — 1996. — т. 166, № 11. — с. 1145—1170.
110. Работа с несбалансированными данными. — URL: <https://developers.google.com/machine-learning/data-prep/construct/sampling-splitting/imbalanced-data?hl=ru> ; (Accessed: 2022).
111. Разбор Precision, Recall, F1 Score и других метрик классификации. — URL: <https://vk.com/@itresume-sobaka-ili-koshka-razbor-precision-recall-i-f1-score-i-drug> ; (Accessed: 2022).

112. Учебник по машинному обучению, Yandex. — URL: <https://academy.yandex.ru/handbook/ml> ; (Accessed: 2022).
113. *Watanabe S.* Tree-Structured Parzen Estimator: Understanding Its Algorithm Components and Their Roles for Better Empirical Performance. — 2023. — arXiv: [2304.11127](https://arxiv.org/abs/2304.11127) [cs.LG].
114. *Hutter F., Hoos H. H., Leyton-Brown K.* Sequential model-based optimization for general algorithm configuration // Learning and Intelligent Optimization: 5th International Conference, LION 5, Rome, Italy, January 17-21, 2011. Selected Papers 5. — Springer. 2011. — с. 507—523.
115. Интерпретация результатов машинного обучения. — URL: <https://webiomed.ru/blog/interpretatsiia-rezultatov-mashinnogo-obucheniia/> ; (Accessed: 2022).
116. *Podoprosvetov A., Alisejchik A., Orlov I.* Comparison of action recognition from video and IMUs // Procedia Computer Science. — 2021. — т. 186. — с. 242—249.
117. Методы управления движением искусственной мышцы и экзоскелета верхней части тела на ее основе / М. А. Чумичев [и др.] // Препринты Института прикладной математики им. МВ Келдыша РАН. — 2021. — № 37. — с. 1—24.
118. Variational Autoencoder / L. Pinheiro Cinelli [и др.] // Variational Methods for Machine Learning with Applications to Deep Networks. — Cham : Springer International Publishing, 2021. — с. 111—149. — ISBN 978-3-030-70679-1. — DOI: [10.1007/978-3-030-70679-1_5](https://doi.org/10.1007/978-3-030-70679-1_5). — URL: https://doi.org/10.1007/978-3-030-70679-1_5.
119. *Gneiting T., Raftery A. E.* Strictly proper scoring rules, prediction, and estimation // Journal of the American statistical Association. — 2007. — т. 102, № 477. — с. 359—378.
120. Generative time series forecasting with diffusion, denoise, and disentanglement / Y. Li [и др.] // Advances in Neural Information Processing Systems. — 2022. — т. 35. — с. 23009—23022.
121. Generative Time Series Forecasting with Diffusion, Denoise, and Disentanglement / Y. Li [и др.]. — 2023. — arXiv: [2301.03028](https://arxiv.org/abs/2301.03028) [cs.LG].

Список рисунков

1.1	Пример работы системы <i>OpenPose</i> на кадрах камер с двух ракурсов: (а) – ракурс 1, (б) – ракурс 2.	38
1.2	(а) Номера ключевых точек, получаемых системой <i>OpenPose</i> . (б) Расположение датчиков IMU на теле человека.	39
1.3	Схема установки камер	40
1.4	Отображение точек скелета человека	40
1.5	3D-модель человека, построенная по кадрам.	45
1.6	Для относительного изменения длины плеч коэффициент выполнения условия оценивается в 0.2 для 5 минутного отрезка времени	45
1.7	График интерполированных точек после фильтрации	47
1.8	Гистограмма коэффициентов уверенности точки 0.	48
2.1	Архитектура нейронной сети с LSTM слоями	56
2.2	Архитектура нейронной сети с CNN слоями	58
2.3	Представление данных в обучающем наборе	60
2.4	Пример работы метода по определению действий по 10 кадрам	62
2.5	Пример работы метода по определению действий по 30 кадрам	62
3.1	Схема представления двух звеньев и сустава	74
3.2	Простейшая модель Хилла мышцы с сухожилием	92
3.3	Результаты расчета мощности метаболизма по стандарту <i>ISO8996</i> для пятиминутного эксперимента	97
3.4	Расположение <i>MEMS</i> сенсоров (синие) на теле человека.	105
3.5	Внешний вид системы визуализации	110
3.6	Иллюстрация работы калибровки и визуализации	111
3.7	СК сегмента и СК <i>MEMS</i> сенсора.	113
3.8	СК тела человека и СК <i>kinect</i>	114
3.9	Проекция на экран	115
4.1	Точность по метрике <i>masco</i> <i>goc-auc</i>	126
4.2	Определение моделью целевой переменной 2 на данных с ракурса 2 и набором признаков 2 по видео.	130

- 4.3 Определение моделью целевой переменной 1 по MEMS данным. . 131
- 4.4 Определение моделью целевой переменной 2 на данных с ракурса 2 и набором признаков 1 по видео данным. 131
- 4.5 Кадр, на котором туловище человека заслоняет руку 132

Список таблиц

1	Общий список действий для классификации	33
2	Средняя оценка правдоподобности звеньев модели	46
3	Баланс выборок для очищенных данных	52
4	Результаты двух метрик на контрольной выборке для шести целевых переменных	54
5	Результаты анализа метода	61
6	Модель скелета	68
7	Модель мышечного скелета	69
8	Скелетно-мышечная модель тела человека	71
9	Метрики оценки модели классификации XGBoost, для каждой подвыборки (1-4)	125
10	Метрики $ROC - AUC$ и $Accuracy$ на контрольных данных для первой целевой переменной. (Поднятие руки)	129
11	Метрики $ROC - AUC$ и $Accuracy$ на контрольных данных для второй целевой переменной (покраска стены).	130
12	Сравнительная таблица на основе метрик MSE и $CRPS$	140
13	Оценки качества данных	165
14	Оценка качества часть 2	166
15	Пример разметки данных	167
16	Оценка качества распознанных скелетных данных	168
17	Оценка качества переходов последовательностей движений	169

Приложение А

Демонстрация результата обработки видео данных эксперимента по сбору последовательностей движений

Таблица 13 — Оценки качества данных

Номер	Папка	Видеозапись	Кол-во кадров	Если разметка	Есть ли файл Json	Оценка
1	K_E_D_	1F	7281	Да	Да	Плохо
2	K_E_D_	1T	9805	Да	Да	Плохо
3	K_E_D_	2F	7326	Да	Да	Хорошо
4	K_E_D_	2T	9758	Да	Да	Плохо
5	K_E_D_	3F	7353	Да	Да	Плохо
6	K_E_D_	3T	9701	Да	Да	Плохо
7	K_E_D_	4F	7332	Да	Да	Нормально
8	K_E_D_	4T	9590	Да	Да	Нормально
9	K_E_D_	5F	7380	Да	Да	Нормально
10	K_E_D_	5T	9472	Да	Да	Плохо
11	K_E_D_	6F	7371	Да	Да	Нормально
12	K_E_D_	6T	10243	Да	Да	Плохо
13	K_E_D_	7F	9144		Да	Нормально
14	K_E_D_	7T	9633		Да	Плохо
15	K_E_D_	8F	9219	Да	Да	Плохо
16	K_E_D_	8T	9613	Да	Да	Плохо
17	K_E_D_	9F	9156	Да	Да	Хорошо
18	K_E_D_	9T	9855	Да	Да	Хорошо
19	K_E_D_	10F	9096	Да	Да	Плохо
20	K_E_D_	10T	9409	Да	Да	Хорошо
21	K_E_D_	11F	9099		Да	Хорошо
22	K_E_D_	11T	9584		Да	Хорошо
23	K_E_D_	12F	9297	Да	Да	Хорошо
24	K_E_D_	12T	9576	Да	Да	Плохо
25	K_E_D_	13F	9066		Да	Хорошо
26	K_E_D_	13T	9501		Да	Плохо
27	K_E_D_	14F	9045		Да	Хорошо
28	K_E_D_	14T	9426		Да	Плохо
29	K_E_D_	15F	9099		Да	Хорошо
30	K_E_D_	15T	9531		Да	Плохо
1	M.C.K.	0002_right	18150	Да	Да	Нормально
2	M.C.K.	0002_left	18406	Да	Да	Нормально
3	M.C.K.	0003_right	18248	Да	Да	Нормально
4	M.C.K.	0003_left	17884	Да	Да	Нормально
5	M.C.K.	0004_right	17852	Да	Да	Хорошо
6	M.C.K.	0004_left	17466	Да	Да	Нормально
7	M.C.K.	0005_right	18182	Да	Да	Плохо
8	M.C.K.	0005_left	18107	Да	Да	Хорошо
9	M.C.K.	0006_right	18333	Да	Да	Хорошо
10	M.C.K.	0006_left	18037	Да	Да	Плохо
11	M.C.K.	0007_right	18320	Да	Да	Плохо
12	M.C.K.	0007_left	18248	Да	Да	Плохо
13	M.C.K.	0008_right	17331	Да	Да	Хорошо
14	M.C.K.	0008_left	17663	Да	Да	Хорошо
15	M.C.K.	0009_right	15862	Да	Да	Плохо
16	M.C.K.	0009_left	17383	Да	Да	Плохо

Таблица 14 — Оценка качества часть 2

Номер	Папка	Видеозапись	Кол-во кадров	Если разметка	Есть ли файл Json	Оценка
1	M_Д_Д	0004_right	388		Да	Плохо
2	M_Д_Д	0004_left	394		Да	Плохо
3	M_Д_Д	0005_right	365		Да	Плохо
4	M_Д_Д	0005_left	365		Да	Плохо
5	M_Д_Д	0006_right	369		Да	Плохо
6	M_Д_Д	0006_left	369		Да	Плохо
7	M_Д_Д	0007_right	21419		Да	Плохо
8	M_Д_Д	0007_left	11231		Да	Плохо
9	M_Д_Д	0008_right	22482		Да	Плохо
10	M_Д_Д	0008_left	10943		Нет	Плохо
11	M_Д_Д	0009_right	21906		Да	Плохо
12	M_Д_Д	0009_left	11926		Да	Плохо
13	M_Д_Д	0010_right	23871		Да	Плохо
14	M_Д_Д	0010_left	10732		Да	Плохо
15	M_Д_Д	0011_right	8597		Да	Плохо
16	M_Д_Д	0011_left	10998		Да	Плохо
17	M_Д_Д	0012_right	11011		Да	Хорошо
18	M_Д_Д	0012_left	10835		Да	Хорошо
19	M_Д_Д	0013_right	10844		Да	Плохо
20	M_Д_Д	0013_left	10487		Да	Нормально
21	M_Д_Д	0014_right	10499		Да	Хорошо
22	M_Д_Д	0014_left	10790		Да	Плохо
23	M_Д_Д	0015_right	10800		Да	Плохо
24	M_Д_Д	0015_left	11010		Да	Плохо
25	M_Д_Д	0016_right	10850		Да	Нормально
26	M_Д_Д	0016_left	10997		Да	Плохо
27	M_Д_Д	0017_right	10860		Да	Нормально
28	M_Д_Д	0017_left	13155		Да	Нормально
29	M_Д_Д	0018_right	13070		Да	Плохо
30	M_Д_Д	0018_left	10920		Да	Плохо
31	M_Д_Д	0019_right	10931		Да	Нормально
32	M_Д_Д	0019_left	11047		Да	Плохо
33	M_Д_Д	0020_right	10023		Да	Плохо
34	M_Д_Д	0020_left	11161		Да	Плохо
35	M_Д_Д	0021_right	10564		Да	Нормально
36	M_Д_Д	0021_left	10967		Да	Плохо
37	M_Д_Д	0022_right	10969		Да	Плохо
38	M_Д_Д	0022_left	9794		Да	Плохо
39	M_Д_Д	0023_right	9788		Да	Плохо
40	M_Д_Д	0023_left	13066		Да	Плохо
41	M_Д_Д	0024_right	13061		Да	Плохо
42	M_Д_Д	0024_left	12790		Да	Плохо
43	M_Д_Д	0025_right			Нет	Плохо
44	M_Д_Д	0025_left	12787		Да	Нормально
45	M_Д_Д	0026_right	14435		Да	Нормально
46	M_Д_Д	0026_left			Нет	Плохо
47	M_Д_Д	0027_right	14438		Да	Плохо
48	M_Д_Д	0027_left	15096		Да	Плохо
49	M_Д_Д	0028_right	502		Да	Плохо
50	M_Д_Д	0028_left	503		Да	Плохо
51	M_Д_Д	0029_right	427		Да	Плохо
52	M_Д_Д	0029_left	427		Да	Плохо
53	M_Д_Д	0030_right	395		Да	Плохо
54	M_Д_Д	0030_left	395		Да	Плохо

Далее в таблице приведен пример разметки данных, момента когда человек выполняет действие подъем большого предмета, касается коробки не поднимая ее ставит назад и заканчивает действие.

Таблица 15 — Пример разметки данных

Номер кадра	Имя кадра	zahvat_predmeta	podzem_bolshogo_predmeta
7140	frame_9T_007140.jpg	0	0
7145	frame_9T_007145.jpg	0	0
7146	frame_9T_007146.jpg	0	0
7147	frame_9T_007147.jpg	0	1
7148	frame_9T_007148.jpg	0	1
7149	frame_9T_007149.jpg	0	1
7150	frame_9T_007150.jpg	0	1
7151	frame_9T_007151.jpg	0	1
7152	frame_9T_007152.jpg	0	1
7153	frame_9T_007153.jpg	0	1
7154	frame_9T_007154.jpg	0	1
7155	frame_9T_007155.jpg	0	1
7156	frame_9T_007156.jpg	0	1
7157	frame_9T_007157.jpg	0	1
7158	frame_9T_007158.jpg	0	1
7159	frame_9T_007159.jpg	0	1
7160	frame_9T_007160.jpg	0	1
7161	frame_9T_007161.jpg	0	1
7162	frame_9T_007162.jpg	0	1
7163	frame_9T_007163.jpg	0	1
7164	frame_9T_007164.jpg	0	1
7165	frame_9T_007165.jpg	0	1
7166	frame_9T_007166.jpg	0	1
7167	frame_9T_007167.jpg	0	1
7168	frame_9T_007168.jpg	1	1
7169	frame_9T_007169.jpg	1	1
7170	frame_9T_007170.jpg	1	1
7171	frame_9T_007171.jpg	1	1
7172	frame_9T_007172.jpg	1	1
7173	frame_9T_007173.jpg	1	1
7174	frame_9T_007174.jpg	1	1
7175	frame_9T_007175.jpg	1	1
7176	frame_9T_007176.jpg	1	1
7177	frame_9T_007177.jpg	1	1
7178	frame_9T_007178.jpg	1	1
7179	frame_9T_007179.jpg	1	1
7180	frame_9T_007180.jpg	1	1
7181	frame_9T_007181.jpg	1	1
7182	frame_9T_007182.jpg	1	1
7183	frame_9T_007183.jpg	1	1
7184	frame_9T_007184.jpg	0	1
7185	frame_9T_007185.jpg	0	1
7186	frame_9T_007186.jpg	0	1
7187	frame_9T_007187.jpg	0	1
7188	frame_9T_007188.jpg	0	1
7189	frame_9T_007189.jpg	0	1
7190	frame_9T_007190.jpg	0	1
7191	frame_9T_007191.jpg	0	1

Таблица 16 — Оценка качества распознанных скелетных данных

Номер	Оценка	Описание (секунды - ошибка происходит в первый раз. Повторяется позже ещё раз, но я не указываю)
1	Плохо	Есть разрывы в конечностях, 3:10 рука стала ногой
2	Плохо	Рука становится ногой, скелет ломается, руки меняются местами, проблема с коробкой (ноги)
3	Хорошо	Нога немного ломается
4	Плохо	Проблемы с коробкой (ноги), руки меняются местами
5	Плохо	0:13 нога становится рукой, 1:25 рука переходит в голову
6	Плохо	0:37 нога ломается, 1:07 рука становится ногой
7	Нормально	Руки меняются местами, ноги ломаются
8	Нормально	Руки меняются местами, ноги ломаются
9	Нормально	Скелет ломается периодически, руки меняются местами, ноги ломаются
10	Плохо	Рука становится ногой, руки меняются местами, ноги ломаются
11	Нормально	Фокус видео, ноги исчезают
12	Плохо	Скелет ломается, руки становятся ногами, руки меняются местами
13	Нормально	Нога исчезает, рука становится ногой (однажды), скелет исчезает немного
14	Плохо	Плохой фокус видео данных
15	Плохо	Ноги исчезают, весь скелет исчезает немного
16	Плохо	Руки меняются местами, руки становятся ногами
17	Хорошо	Скелет исчезает (но не при вып. действия)
18	Хорошо	Руки меняются местами
19	Плохо	Ноги исчезают, скелет частично исчезает, нога ломается
20	Хорошо	Рука становится ногой, нога ломается
21	Хорошо	Нога ломается
22	Хорошо	Нога ломается, рука становится ногой
23	Хорошо	Немного ломается
24	Плохо	Плохое качество видео
25	Хорошо	Нога ломается, рука становится ногой
26	Плохо	Нога ломается
27	Хорошо	Нога ломается
28	Плохо	Нога ломается, скелет очень ломается
29	Хорошо	Нога ломается, исчезновение ноги, ноги меняются местами
30	Плохо	Нога ломается, рука становится ногой, скелет исчезает вообще, рука становится ногой
1	Нормально	Скелет ломается, когда поднимает коробку
2	Нормально	Скелет ломается, когда поднимает коробку
3	Нормально	Скелет частично исчезает при подборе предмета, ноги меняются местами, немного ломаются
4	Нормально	Одна из ног становится ножкой табурета, ноги меняются местами, немного ломаются
5	Хорошо	
6	Нормально	Скелет не на человеке, а на картине!
7	Плохо	Скелет ломается, рука становится ногой, логи ломаются
8	Хорошо	
9	Хорошо	Логги ломаются (дважды) при нагибании, скелет исчезает, нога становится частью коробки
10	Плохо	Логги ломаются
11	Плохо	Скелет исчезает частично, руки становятся ногами, ноги исчезают
12	Плохо	При наклоне исчезает скелет, скелет ломается, скелет не на человеке, а на картине
13	Хорошо	Скелет не на человеке ненадолго
14	Хорошо	
15	Плохо	Скелета почти нет!
16	Плохо	
1	Плохо	Руки становятся ногами, скелет исчезает, скелет на стене, скелет идет на потолок
2	Плохо	Скелет исчезает, нет ног, руки становятся ногами, скелет на полу
3	Плохо	Рука становится кашпоном, ноги пропадают, скелет исчезает
4	Плохо	Скелет пропадает, скелет на полу, ноги пропадают
5	Плохо	Скелет омается, ноги переходят на стену, пропадают ноги, пропадает сам скелет
6	Плохо	Нагибается и скелет пропадает, скелет на полу, скелет исчезает, ноги исчезают
7	Плохо	Скелет исчезает, нет ног, нагибается и нет скелета, исчезает скелет
8	Плохо	Скелет исчезает, скелет на полу, рука переходит на тень этой руки на стене

Таблица 17 — Оценка качества переходов последовательностей движений

Номер	Оценка переходов	комментарий по переходам
1	Нормально	действия выполняются подряд (пауза на 0:55-1:03, 1:12-1:18, 2:40-2:49, 2:56-3:04, 3:12-3:19)
3	Плохо	Очень длинные паузы
5	Плохо	Очень длинные паузы
7	Плохо	Очень длинные паузы
9	Плохо	Очень длинные паузы
11	Плохо	Очень длинные паузы
13	Нормально	паузы короче и их меньше
15	Нормально	Аналогично предыдущему
17	Плохо	Очень длинные паузы
1	Хорошо	Длинные действия, но переходы четкие и без пауз
3	Хорошо	Длинные действия, но переходы четкие и без пауз
5	Хорошо	Переходов мало, но они четкие и естественные
7	Хорошо	Переходы медленнее, но тоже естественные
9	Хорошо	Переходы быстрее чем в предыдущем, нормальные
11	Хорошо	Меньше переходов, но они по-прежнему естественные
13	Плохо	Одно действие идет все видео
15	Нормально	Мало переходов
1	Нормально	Видео right слишком темные до 0010
3	Нормально	Мало переходов
5	Нормально	Мало переходов
7	Нормально	В середине видео 2:55-3:03 пауза
9	Плохо	Между первым действием на 5 секунде и вторым на 0:55 ничего не происходит
11	Плохо?	С 4:40 до конца просто стоит
13	Нормально	Переход с почесыванием головы выглядит не совсем естественно
15	Плохо	Длинные переходы, паузы в действиях
17	Плохо	1 действие все видео, при этом не заканчивается четко
19	Плохо	2:20-2:40 стоит 2:40-3:20 работа в телефоне, между действиями далее пауза
21	Плохо	4:24-4:54 пауза 30 секунд
23	Плохо	пауза 2:05-2:27
25	Плохо	2:49-3:00 неестественная пауза, с 5:14 до конца стоит
27	Плохо	Прерывается на объяснение действий, долгая пауза
29	Плохо	Стоит по 10-20 секунд между действиями
31	Плохо	Длинные паузы
33	Плохо	Одно действие с длинной паузой
35	Плохо	Аналогично с прошлым
37	Нормально?	Работа в телефоне 3:18-4:28
39	Плохо	Длинные паузы между действиями
41	Нормально?	Телефон 3:06-4:34
45	Нормально	Всего два действия 25 лефт и 26 райт одно видео
47	Нормально?	3:10-5:00 телефон
49	Плохо	По сути нет переходов кроме начального, одно действие
51	Плохо	2:28-3:50 стоит 3:50-4:20 работа в телефоне
53	Плохо	По сути нет переходов кроме начального, одно действие