

Министерство науки и высшего образования Российской Федерации
 федеральное государственное автономное
 образовательное учреждение высшего образования
 «Национальный исследовательский Томский политехнический университет» (ТПУ)

Инженерная школа информационных технологий и робототехники
 Направление подготовки 09.04.01 «Информатика и вычислительная техника»
 Отделение информационных технологий

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Тема работы
Разработка алгоритма автоматического восстановления позы скелета руки на основе карт глубины

УДК 004.514:611.717:004.932.72'1

Студент

Группа	ФИО	Подпись	Дата
8BM73	Иовлев Юлий Денисович		

Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ОИТ ИШИТР	Спицын Владимир Григорьевич	д.т.н., профессор		

КОНСУЛЬТАНТЫ ПО РАЗДЕЛАМ:

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОСГН	Меньшикова Екатерина Валентиновна	к.ф.н., доцент		

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Ассистент ООД	Алексеев Николай Архипович	—		

ДОПУСТИТЬ К ЗАЩИТЕ:

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
09.04.01 "Информатика и вычислительная техника", профиль "Компьютерный анализ и интерпретация данных"	Спицын Владимир Григорьевич	д.т.н., профессор		

Томск – 2019 г.

**ЗАПЛАНИРОВАННЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ОСНОВНОЙ
ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЕ ПОДГОТОВКИ МАГИСТРОВ 09.04.01
«ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА», ИШИТР ТПУ, ПРОФИЛЬ
«КОМПЬЮТЕРНЫЙ АНАЛИЗ И ИНТЕРПРЕТАЦИЯ ДАННЫХ»**

Код результата	Результат обучения (выпускник должен быть готов)
P1	Применять базовые и специальные естественнонаучные и математические знания в области информатики и вычислительной техники, достаточные для комплексной инженерной деятельности.
P2	Применять базовые и специальные знания в области современных информационных технологий для решения инженерных задач.
P3	Ставить и решать задачи комплексного анализа, связанные с созданием аппаратно-программных средств информационных и автоматизированных систем, с использованием базовых и специальных знаний, современных аналитических методов и моделей.
P4	Разрабатывать программные и аппаратные средства (системы, устройства, блоки, программы, базы данных и т. п.) в соответствии с техническим заданием и с использованием средств автоматизации проектирования.
P5	Проводить теоретические и экспериментальные исследования, включающие поиск и изучение необходимой научно-технической информации, математическое моделирование, проведение эксперимента, анализ и интерпретация полученных данных, в области создания аппаратных и программных средств информационных и автоматизированных систем.
P6	Внедрять, эксплуатировать и обслуживать современные программно-аппаратные комплексы, обеспечивать их высокую эффективность, соблюдать правила охраны здоровья, безопасность труда, выполнять требования по защите окружающей среды.
Универсальные компетенции	
P7	Использовать базовые и специальные знания в области проектного менеджмента для ведения комплексной инженерной деятельности.
P8	Владеть иностранным языком на уровне, позволяющем работать в иноязычной среде, разрабатывать документацию, презентовать и защищать результаты комплексной инженерной деятельности.
P9	Эффективно работать индивидуально и в качестве члена группы, состоящей из специалистов различных направлений и квалификаций, демонстрировать ответственность за результаты работы и готовность следовать корпоративной культуре организации.
P10	Демонстрировать знания правовых, социальных, экономических и культурных аспектов комплексной инженерной деятельности.
P11	Демонстрировать способность к самостоятельному обучению в течение всей жизни и непрерывному самосовершенствованию в инженерной профессии.

Министерство науки и высшего образования Российской Федерации
федеральное государственное автономное
образовательное учреждение высшего образования
«Национальный исследовательский Томский политехнический университет» (ТПУ)

Инженерная школа информационных технологий и робототехники
Направление подготовки 09.04.01 «Информатика и вычислительная техника»
Отделение информационных технологий

УТВЕРЖДАЮ:
Руководитель ООП

(Подпись) (Дата) (Ф.И.О.)

ЗАДАНИЕ на выполнение выпускной квалификационной работы

В форме:

Магистерской диссертации

Студенту:

Группа	ФИО
8BM73	Иовлеву Юлию Денисовичу

Тема работы:

Разработка алгоритма автоматического восстановления позы скелета руки на основе карт глубины	
Утверждена приказом директора (дата, номер)	№ 1290/с 19.02.2019

Срок сдачи студентом выполненной работы:

27.05.2019

ТЕХНИЧЕСКОЕ ЗАДАНИЕ:

Исходные данные к работе	Набор изображений, полученных с датчиков глубины.
Перечень подлежащих исследованию, проектированию и разработке вопросов	– обзор предметной области, анализ научной литературы; – изучение общедоступных методов, применимых для детектирования положения руки; – разработка собственных алгоритмов детектирования положения руки; – программная реализация предложенных алгоритмов;

	– оценка эффективности и практической применимости разработанных алгоритмов; – финансовый менеджмент; - ресурсоэффективность и ресурсосбережение; – социальная ответственность.
Перечень графического материала	– общая схема работы нейросетевого алгоритма; – UML диаграмма развертывания системы; – UML диаграмма компонентов.
Консультанты по разделам выпускной квалификационной работы	
Раздел	Консультант
Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	Меньшикова Екатерина Валентиновна
Социальная ответственность	Алексеев Николай Архипович
Названия разделов, которые должны быть написаны на русском и иностранном языках:	
Анализ предметной области	

Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику	
---	--

Задание выдал руководитель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ОИТ ИШИТР	Спицын Владимир Григорьевич	д.т.н., профессор		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8BM73	Иовлев Юлий Денисович		

Министерство науки и высшего образования Российской Федерации
 федеральное государственное автономное
 образовательное учреждение высшего образования
 «Национальный исследовательский Томский политехнический университет» (ТПУ)

Инженерная школа информационных технологий и робототехники
 Направление подготовки 09.04.01 «Информатика и вычислительная техника»
 Уровень образования магистратура
 Отделение информационных технологий
 Период выполнения осенний / весенний семестр 2018/2019 учебного года

Форма представления работы:

Магистерская диссертация

КАЛЕНДАРНЫЙ РЕЙТИНГ-ПЛАН
выполнения выпускной квалификационной работы

Срок сдачи студентом выполненной работы:	27.05.2019
--	------------

Дата контроля	Название раздела (модуля) / вид работы (исследования)	Максимальный балл раздела (модуля)
25.10.18	Аналитический обзор	15
25.01.19	Разработка алгоритмов для детектирования положения руки	30
28.03.19	Программная реализация разработанных алгоритмов	20
29.04.19	Тестирование реализованных программных средств	15
17.05.19	Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	10
17.05.19	Социальная ответственность	10

СОСТАВИЛ:

Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ОИТ ИШИТР	Спицын Владимир Григорьевич	д.т.н., профессор		

СОГЛАСОВАНО:

Руководитель ООП

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ОИТ ИШИТР	Спицын Владимир Григорьевич	д.т.н., профессор		

ЗАДАНИЕ ДЛЯ РАЗДЕЛА «ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И РЕСУРСОСБЕРЕЖЕНИЕ»

Студенту:

Группа	ФИО
8BM73	Иовлев Юлий Денисович

Школа	Информационных технологий и робототехники	Отделение школы (НОЦ)	Отделение информационных технологий
Уровень образования	Магистратура	Направление/специальность	09.04.01 Информатика и вычислительная техника

Исходные данные к разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»:

1. Стоимость ресурсов научного исследования (НИ): материально-технических, энергетических, финансовых, информационных и человеческих	Оклад руководителя – 47104 руб. Оклад инженера – 12663 руб Общая сумма затрат – 389379 руб.
2. Нормы и нормативы расходования ресурсов	Тариф на электроэнергию – 5,8 руб/кВт Норма амортизации – 33%
3. Используемая система налогообложения, ставки налогов, отчислений, дисконтирования и кредитования	Ставка для расчета отчислений во внебюджетные фонды – 27,1% Ставка для расчета накладных расходов – 30%

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

1. Оценка коммерческого и инновационного потенциала НТИ	Определение потенциальных потребителей, анализ конкурентных решений с позиции ресурсоэффективности и экономической эффективности, функционально-стоимостной анализ.
2. Разработка устава научно-технического проекта	Цели и ожидаемые результаты исследования. Трудозатраты и основные функции исполнителей проекта.
3. Планирование процесса управления НТИ: структура и график проведения, бюджет, риски и организация закупок	Составление перечня этапов и работ, расчёт затрат на необходимые ресурсы, описание возможных рисков.
4. Определение ресурсной, финансовой, экономической эффективности	Оценка интегрального показателя эффективности НИР, при помощи расчёта финансовой и ресурсной эффективности.

Перечень графического материала:

1. «Портрет» потребителя результатов НТИ
2. Сегментирование рынка
3. Оценка конкурентоспособности технических решений
4. Матрица SWOT
5. График проведения и бюджет НТИ
6. Оценка ресурсной, финансовой и экономической эффективности НТИ
7. Потенциальные риски

Дата выдачи задания для раздела по линейному графику

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОСТН	Меньшикова Е.В	к.ф.н, доцент		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8BM73	Иовлев Ю.Д.		

ЗАДАНИЕ ДЛЯ РАЗДЕЛА «СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»

Студенту:

Группа	ФИО
8BM73	Иовлев Юлий Денисович

Школа	Информационных технологий и робототехники	Отделение (НОЦ)	Отделение информационных технологий
Уровень образования	Магистратура	Направление/специальность	09.04.01 Информатика и вычислительная техника

Исходные данные к разделу «Социальная ответственность»:	
Характеристика объекта исследования (вещество, материал, рабочая зона) и области его применения	Разработка алгоритма автоматического восстановления позы скелета руки на основе карт глубины
Перечень вопросов, подлежащих исследованию, проектированию и разработке:	
1. Правовые и организационные вопросы обеспечения безопасности	1. Специальные правовые нормы трудового законодательства; 2. Организационные мероприятия при компоновке рабочей зоны.
2. Производственная безопасность 2.1. Анализ выявленных вредных факторов при разработке и эксплуатации проектируемого решения 2.2. Анализ выявленных опасных факторов при разработке и эксплуатации проектируемого решения	Вредные факторы: 1. Отклонения показателей микроклимата; 2. Недостаточная освещенность рабочей зоны; 3. Превышение уровня шума; 4. Повышенный уровень электромагнитных излучений; Опасные факторы: 1. Электрический ток 2. Опасность возникновения пожара
3. Экологическая безопасность	Источники выбросов в атмосферу; Образование сточных вод и отходов. Мероприятия по снижению вредного воздействия на ОС
4. Безопасность в чрезвычайных ситуациях	Вероятные ЧС при разработке и эксплуатации проектируемого решения и меры по их предупреждению

Дата выдачи задания для раздела по линейному графику	
--	--

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Ассистент ООД ШБИП	Алексеев Н.А.	—		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8BM73	Иовлев Ю.Д.		

РЕФЕРАТ

Выпускная квалификационная работа содержит 104 страниц, 27 рисунка, 30 таблиц, 35 источника и 2 приложения.

Ключевые слова: оценка позы, машинное обучение, глубокое обучение, карта глубины, облако точек, компьютерная геометрия.

Объектом исследования являются карты глубины, полученные с датчиков. Предметом исследования являются алгоритмы обработки данных с датчиков.

Цель работы: разработка алгоритма автоматического восстановления позы скелета руки на основе карт глубины.

В процессе исследования были изучены и проанализированы существующие методы и подходы для детектирования позы скелета руки на основе карт глубины.

В результате исследования были предложены и программно-реализованы два подхода: алгоритмический подход к детектированию позы руки, а также подход с использованием сверточных нейронных сетей.

Основные конструктивные, технологические и технико-эксплуатационные характеристики: программное приложение, позволяющее выполнять детектирование позы руки человека на основе данных с датчиков глубины.

Степень внедрения: используется в компании ООО «Рубиус Групп».

Область применения: детектирование и уточнение положения суставов скелета руки в трехмерном пространстве.

Значимость работы заключается в практической разработке методов обработки карт глубины, позволяющих с высокой точностью детектировать положения суставов руки в трехмерном пространстве.

Запланированным развитием работы является оптимизация предложенных алгоритмов с целью увеличения их производительности, а также повышение качества детектирования.

ОПРЕДЕЛЕНИЯ, ОБОЗНАЧЕНИЯ, СОКРАЩЕНИЯ И НОРМАТИВНЫЕ ССЫЛКИ

2D, 3D (2-dimensional, 3-dimensional) – что-либо имеющее два измерения, три измерения.

API (application programming interface) – Набор готовых классов, процедур, функций, структур и констант, предоставляемых приложением (библиотекой, сервисом) для использования во внешних программных продуктах.

ICP (iterative closest point) – итеративный алгоритм ближайших точек.

MSE (mean square error) – средний квадрат ошибки при вычислении какой-либо величины.

OBB (oriented bounding box) – тип ограничивающего прямоугольника, используемый в компьютерной геометрии.

CNN (convolutional neural network) – свёрточная нейронная сеть, одна из современных архитектур искусственных нейронных сетей.

SGD (stochastic gradient descent) – стохастический градиентный спуск, итеративный метод оптимизации дифференцируемой целевой функции.

LCN (local contrast normalization) – метод, используемый для нормализации контраста изображений.

Оглавление

ВВЕДЕНИЕ	13
1 Анализ предметной области	15
1.1 Описание предметной области	15
1.2 Обзор литературы	15
1.3 Теоретические основы	17
1.3.1 Карты глубины	17
1.3.2 Матрица камеры	18
1.3.3 Скелетная анимация	20
2 РАЗРАБОТКА АЛГОРИТМОВ восстановления позы скелета руки	22
2.1 Методы и средства разработки	22
2.1.1 Языки разработки	22
2.1.2 Среда разработки	22
2.1.3 Используемые средства и фреймворки	23
2.2 Проектирование архитектуры веб-приложения	24
2.2.1 Преимущества и недостатки микросервисной архитектуры	27
2.3 Получение облака точек и пространственная фильтрация	28
2.4 Итеративный алгоритм ближайших точек	29
2.5 Алгоритмический подход к решению задачи	30
2.6 Нейросетевой подход к решению задачи	33
2.6.1 Формализация задачи	36
2.6.2 Проецирование облака точек	38
2.6.3 Архитектура CNN	40
2.6.4 Восстановление положения суставов по тепловым картам	41
2.6.5 Обучение CNN	43

2.6.6	Наборы данных и метрики.....	44
2.7	Тестирование	45
3	РЕЗУЛЬТАТЫ РАБОТЫ.....	47
4	ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И РЕСУРСОСБЕРЕЖЕНИЕ.....	50
4.1	Потенциальные потребители результатов исследования	50
4.2	Сегментирование рынка.....	50
4.3	Оценка конкурентоспособности технических решений.....	51
4.4	SWOT-анализ.....	52
4.5	Оценка готовности проекта к коммерциализации.....	54
4.5.1	Методы коммерциализации результатов научно-технического исследования	56
4.6	Инициация проекта.....	56
4.6.1	Цели и результат проекта.....	56
4.6.2	Организационная структура проекта.	58
4.6.3	Ограничения и допущения проекта.	59
4.7	График проведения и бюджет НТИ	59
4.7.1	Бюджет НТИ.....	63
4.8	Реестр рисков проекта	68
4.9	Оценка ресурсной, финансовой и экономической эффективности НТИ	69
5	СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ.....	72
5.1	Производственная безопасность	72
5.2	Анализ выявленных вредных факторов при разработке и эксплуатации проектируемого решения	73
5.3	Анализ выявленных опасных факторов при разработке и эксплуатации проектируемого решения	81

5.3.1	Электробезопасность.....	81
5.3.2	Пожаробезопасность.....	83
5.4	Экологическая безопасность	83
5.4.1	Анализ влияния процесса исследования на окружающую среду	83
5.4.2	Обоснование мероприятий по защите окружающей среды	84
5.5	Защита в чрезвычайных ситуациях.....	85
5.5.1	Анализ вероятных ЧС, которые могут возникнуть на рабочем месте при проведении исследований.....	85
5.5.2	Обоснование мероприятий по предотвращению ЧС и разработка порядка действия в случае возникновения ЧС	85
5.6	Правовые и организационные вопросы обеспечения безопасности	87
5.6.1	Правовые нормы трудового законодательства	87
5.6.2	Организационные мероприятия при компоновке рабочей зоны исследователя	88
	ЗАКЛЮЧЕНИЕ	92
	СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ	93
	ПРИЛОЖЕНИЕ А (справочное).....	96
	ПРИЛОЖЕНИЕ Б CD-ДИСК.....	104

ВВЕДЕНИЕ

Жесты являются частью нашей повседневной жизни, например, некоторые люди жестикулируют, пока они говорят, чтобы подчеркнуть их намерения, маленькие дети указывают на то, что они хотели бы, люди приветствуют друг друга, просто размахивая руками.

Таким образом, жесты являются привычной, интуитивно-понятной и простой формой коммуникации, а задача оценки человеческой позы и выявление жестов вызывает всё больше интереса в области компьютерного зрения.

Жесты могут использоваться для бесконтактного управления приложением или для игры на консолях, например, Xbox 360 в сочетании с недорогим RGB-D датчиком Microsoft Kinect, который может генерировать 3D облако точек сцены перед камерой. В таком случае подаваемые команды определяются на основе движений или расположений частей тела.

Обнаружение жестов тела и рук требуют эффективных алгоритмов для оценки и классификации всего тела или жестов.

Одна из основных проблем при оценке жестов – это огромное число различных состояний. Например, рука представляет собой очень подвижный объект, имеющий не менее 26 степеней свободы, что даже с учетом ограничений на каждую степень свободы превращается во множество различных жестов.

В связи с вышеизложенным, высокую степень актуальности имеет задача разработки алгоритмов обработки данных с датчиков глубины, детектирования и распознавания различных объектов в облаке точек. Предметом исследования являются алгоритмы обработки данных с датчиков, а объектом являются карты глубины, полученные непосредственно с помощью датчиков.

Целью работы является разработка алгоритма автоматического восстановления позы скелета руки на основе карт глубины. В работе рассматриваются как алгоритмические способы решения поставленной задачи,

так и нейросетевые методы искусственного интеллекта. К области практического применения данной разработки можно отнести реализацию пользовательских интерфейсов для устройств виртуальной и дополненной реальности, создание качественного человеко-компьютерного интерфейса, в частности, для людей с ограниченными возможностями.

1 АНАЛИЗ ПРЕДМЕТНОЙ ОБЛАСТИ

1.1 Описание предметной области

Оценка жеста на основе визуальной информации продемонстрировала значительный прогресс в последние годы.

Повышенный интерес к этой области можно отнести к двум доминирующим факторам: глубокая визуализация и глубокое обучение. Прежде всего, методы оценки жеста руки полностью изменились с появлением датчиков глубины, поскольку такие продукты как MS Kinect и Intel Realsense стали широко распространены и доступны.

Как источник информации 2.5D, глубина значительно разрешает большую часть двусмысленностей, присутствующих в монокуляре RGB. Во-вторых, глубокое обучение кардинально изменило подходы к решению проблем компьютерного зрения.

Проблема оценки позы руки в 3D привлекает много внимания сообщества специалистов компьютерного зрения, поскольку она играет важную роль во взаимодействии человека с компьютером, например, в приложениях виртуальной и дополненной реальности. Несмотря на недавний прогресс в этой области, надежная и точная оценка позы руки остается сложной задачей.

Методы, основанные на обучении некоторой модели на большом объеме данных, такие как самоорганизующиеся карты, случайные леса и сверточные нейронные сети (CNN), набирают популярность. С появлением больших аннотированных наборов данных поз рук, данные подходы становятся более выгодными, поскольку они не требуют сложной калибровки модели и устойчивы к плохой инициализации.

1.2 Обзор литературы

Оценка позы на основе компьютерного зрения широко изучалась в литературе на протяжении многих лет. Наиболее распространенные подходы к оценке позы руки можно классифицировать на два вида: подходы, основанные на моделях, и подходы, основанные на данных [3].

Подходы, основанные на моделях, обычно находят оптимальные параметры позы руки путем подбора деформируемой трехмерной модели руки. Такие методы продемонстрировали свою эффективность, особенно после появления карт глубины [4]. Однако большинство данных подходов обладают недостатками, которые будут рассмотрены в дальнейшем.

Подходы, основанные на данных, не нуждаются в явной спецификации размера руки и ограничений движения. Скорее такая информация автоматически кодируется в обучающих данных. Поэтому многие современные методы построены по такой схеме [5]. Среди них случайный лес, его использование оказалось достаточно точным и быстрым. В статье [6] авторы предлагают использовать случайный лес для прямой регрессии углов суставов на основе карт глубины. Аналогичный метод представлен в [7], в котором дополнительно применяется трансферное обучение, чтобы компенсировать несоответствие между синтезом и данными реального мира. Поскольку предсказания случайного леса могут быть неоднозначными для сложных поз руки, иногда используются заранее изученные априорные позы рук, чтобы лучше объединять независимо предсказанные положения суставов [8].

Совсем недавно сверточные нейронные сети показали свою эффективность в оценке сочлененных поз. В [9] сети настроены на регрессию для 2D-позы человека путем непосредственной минимизации ошибки оценки позы на тренировочных данных. Результаты показали, что в значительной степени превосходят традиционные методы. Однако обучение сети занимает более двадцати дней, а набор данных содержит только несколько тысяч изображений. Учитывая относительно небольшой размер набора данных, использованного в [9], можно предположить невозможность использовать данную модель на больших наборах данных, таких как [10], которые состоят из более чем 70000 изображений.

Кроме того, в статье [11] утверждается, что такое прямое отображение данных сверточной сети в непрерывные 2D или 3D вектора имеет высокую

нелинейность и сложность. Также такие сети имеют низкую способность к обобщению, что затрудняет обучение CNN таким способом.

В работе по оценке позы тела человека [12], CNN используются для прогнозирования тепловых карт положений суставов вместо оценки положений костей в пространстве, и на каждой тепловой карте интенсивность пикселя указывает вероятность нахождения сустава в этой точке. Во время обучения ошибка регрессии определяется, как L2-норма разности между генерируемой и истинной тепловой картой. Таким образом, сеть может быть эффективно обучена, и такие сети достигают самых современных характеристик.

Точно такая же структура была применена в работе по оценке 3D-позы руки [13]. Однако тепловая карта предоставляет только 2D информацию о положениях суставов руки, а информация о глубине используется не полностью. Такие подходы, основанные на тепловых картах, интересны, поскольку тепловые карты могут отражать распределение вероятностей положений трехмерных суставов руки в плоскости проекции.

На основе изученной литературы и работ в данной сфере был сформирован немного иной подход к оценке позы руки. В данной работе происходит генерация тепловых карт нескольких проекций и их дальнейшее объединение, с целью получения конечного распределения вероятностей положения суставов руки в трехмерном пространстве.

1.3 Теоретические основы

1.3.1 Карты глубины

Карта глубины – это вид изображения, который состоит из серых пикселей, определенных значениями от 0 до 255 (рис. 1).

Значение «0» серых пикселей означает, что пиксели расположены в самом отдаленном месте 3D-сцены, в то время как значение «255» серых пикселей означает, что «3D» пиксели расположены в самом ближнем месте. На карте глубины каждый пиксель глубины будет определять положение на оси Z, где будет расположен соответствующий ему 2D-пиксель.

С появлением камер с датчиком глубины, данные карты набирают популярность. Карты глубины используются при решении следующих задач:

- генерация стереограмм, а также создания иллюзии трехмерного пространства;
- в компьютерном зрении для моделирования трехмерных фигур или их реконструкции;
- в компьютерном зрении, чтобы позволить трехмерным изображениям обрабатываться инструментами 2D-изображений.



Рисунок 1 – Карта глубины

1.3.2 Матрица камеры

В компьютерном зрении матрица камеры или проекции представляет собой матрицу размера 3×4 , которая описывает отображение камеры-обскуры от 3D точек в мировом пространстве до 2D точек на изображении.

Матрица камеры получается путем выполнения специальной калибровки камеры. Существуют инструменты, позволяющие выполнять калибровку камеры как в ручном режиме, так и в автоматическом.

3D и 2D координаты изображения связываются следующим отображением $x = P * X$, где x – 2D координаты изображения, P – матрица камеры, X – мировые координаты точки. На рисунке 2 показано геометрическое представление данного преобразования.

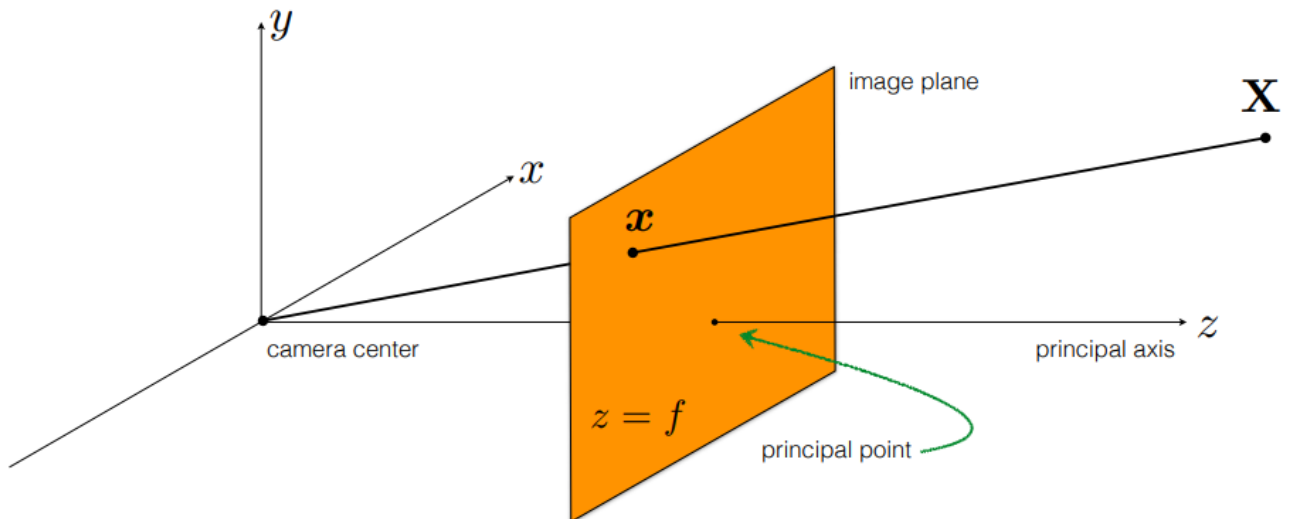


Рисунок 2 – Геометрическая интерпретация

Из рисунка видно, что для выполнения координатных преобразований, необходимо знать фокальное расстояние камеры и основную точку – пересечение главной оси камеры с плоскостью изображения. Данные параметры определяются на этапе калибровки.

Переходя к однородным координатам, получаем следующее соотношение.

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} f & 0 & p_x & 0 \\ 0 & f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1)$$

Полученная формула верна, только если начало мировой системы координат совпадает с положением камеры. В случае когда камеры не находится в начале мировой системы координат, требуется определить смещение и поворот относительно мировой системы координат. Такая матрица называется внешними параметрами камеры и имеет следующий вид.

$$\begin{bmatrix} r1 & r2 & r3 & : & t1 \\ r4 & r5 & r6 & : & t2 \\ r7 & r8 & r9 & : & t3 \end{bmatrix}$$

Левая часть матрицы отвечает за поворот, а правая за смещение.

Таким образом, общая матрица камеры имеет следующий вид.

$$P = \begin{bmatrix} f & 0 & p_x & 0 \\ 0 & f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r1 & r2 & r3 & : & t1 \\ r4 & r5 & r6 & : & t2 \\ r7 & r8 & r9 & : & t3 \end{bmatrix} \quad (2)$$

Имея данную матрицу, можно выполнять преобразования из 2D координат в 3D и обратно.

1.3.3 Скелетная анимация

Скелетная анимация – способ анимирования трёхмерных моделей.

Скелет представляет собой, как правило, древовидную структуру костей, в которой каждая последующая кость связана с предыдущей, то есть наследует все аффинные преобразования с учётом иерархии в скелете [1].

Далее все вершины трехмерной модели соединяются с каким-либо суставом скелета с некоторым коэффициентом. Таким образом, при повороте отдельного сустава происходит движение и всех вершин трехмерной модели, привязанных к нему. Благодаря такому подходу задача скелетной анимации сильно упрощается, за счёт того, что отпадает необходимость анимировать каждую вершину модели в отдельности, а достаточно лишь выполнять аффинные преобразования над суставами скелета.

Шарнирным соединением называется один из основных типов соединений между суставами скелета, такое соединение часто используется при построении скелетных моделей.

В шарнирных соединениях конечная точка одной кости вращается вокруг какой-то статической точки, обычно конечной точки родительской кости. Особенностью такой связи является то, что для перемещения каждой кости в новую позицию задаётся только вращение. Отсюда следует, что для анимации необходимо хранить только информацию о вращении.

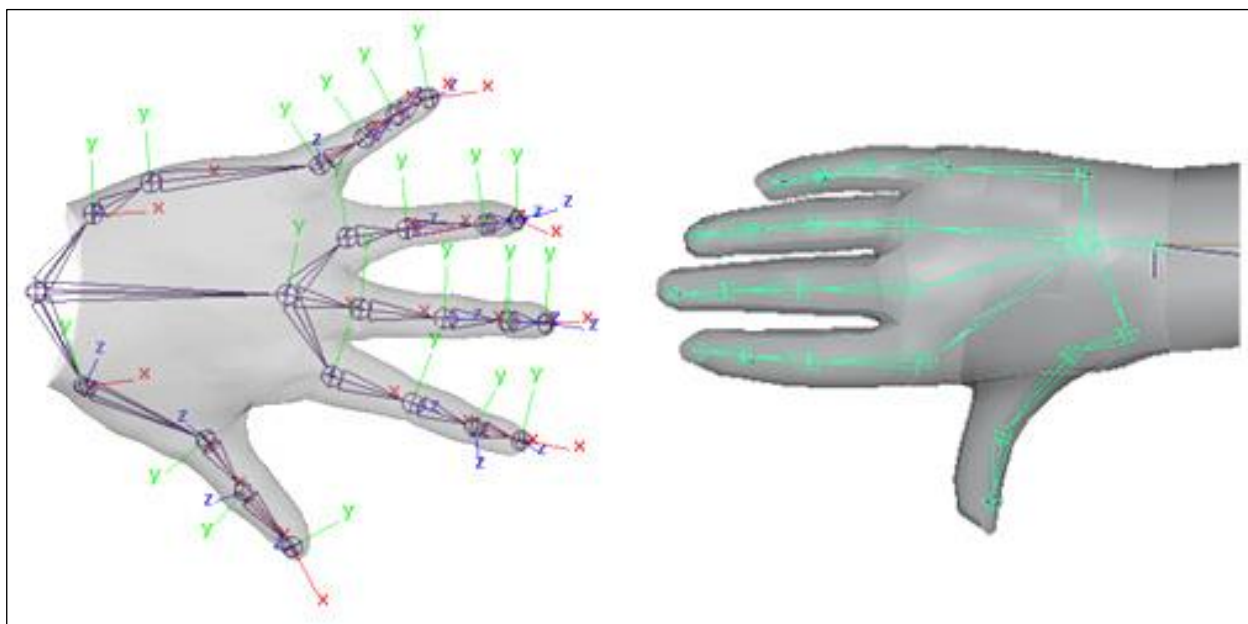


Рисунок 3 – Скелетная модель руки

Таким образом, можно осуществить простой способ задания скелета, на рисунке 3 слева представлен скелет руки с направлением осей для каждого сустава, а справа скелет, связанный с трехмерной моделью. Множество костей имеет иерархическую структуру, каждая кость задается локальным набором параметров – длина и поворот. Поворот обычно задаётся кватернионом или вектором углов поворота.

По договорённости, каждая кость выровнена по оси X. Конечная точка кости находится на оси X и смещена по ней на длину кости. Относительная матрица вращения вычисляется из локального значения вращения текущей кости, а перемещение является сдвигом на длину родительской кости по X-координате в локальной системе координат родителя.

2 РАЗРАБОТКА АЛГОРИТМОВ ВОССТАНОВЛЕНИЯ ПОЗЫ СКЕЛЕТА РУКИ

2.1 Методы и средства разработки

2.1.1 Языки разработки

В качестве языка программирования для серверной части веб-приложения использовалась программная платформа Node.js, основанная на движке V8 (транслирующем JavaScript в машинный код), превращающая JavaScript из узкоспециализированного языка в язык общего назначения [2].

Node.js добавляет возможность JavaScript взаимодействовать с устройствами ввода-вывода через свой API (написанный на C++), подключать другие внешние библиотеки, написанные на разных языках, обеспечивая вызовы к ним из JavaScript-кода. В основе Node.js лежит событийно-ориентированное и асинхронное программирование с неблокирующим вводом/выводом.

Главным преимуществом Node.js является его производительность. Осуществляется это благодаря тому, что используется не стандартная архитектура общения клиента с сервером: для этого языка нет необходимости устанавливать и настраивать веб-сервер.

Для разработки клиентской части веб-приложения использовался язык HTML5, а также каскадные таблицы стилей – CSS.

HTML5 – это новая открытая платформа, предназначенная для создания веб-приложений использующих аудио, видео, графику, анимацию и многое другое.

2.1.2 Среда разработки

Разработка велась на операционной системе Ubuntu 16.04.

Для создания и проектирования приложения в качестве редактора разработки использовалось решение от компании Microsoft – Visual Studio Code, которое обладает мощным функционалом для разработки.

Выбранный редактор подходит одновременно для front-end-разработки и для разработки серверной части на Node.js. Главное преимущество VS Code – это его кроссплатформенность.

Также можно отменить некоторые ключевые возможности, помогающие разработчику:

- отладчик кода;
- интеграция с системами управления версиями;
- огромное количество сторонних расширений.

2.1.3 Используемые средства и фреймворки

Фреймворки являются неотъемлемой частью программного обеспечения, позволяющей упростить и ускорить процесс разработки. Фреймворк – это набор всевозможных библиотек или инструментов для быстрой разработки рутинных задач.

Для реализации микросервиса использовался фреймворк Коа. Коа – это веб-инфраструктура, разработанная командой Express, которая направлена на более выразительную и надежную основу для веб-приложений и API. Используя асинхронные функции, Коа позволяет выполнять обратные вызовы и значительно облегчить обработку ошибок.

На клиентской стороне приложения, для создания нестандартных элементов интерфейса и адаптивной верстки использовались такие популярные фреймворки, как:

- jQuery;
- Bootstrap.

Для создания и отображения анимированной компьютерной 3D графики в браузере использовалась кроссбраузерная JavaScript библиотека Three.js.

Three.js позволяет создавать ускоренную на GPU 3D графику. Это возможно благодаря использованию технологии WebGL.

Основным достоинством при выборе данной библиотеки была возможность импорта трёхмерных моделей формата Collada.

Для совмещения двух облаков точек и нахождения наилучшего их наложения использовалась библиотека libpointmatcher. Данная библиотека основана на алгоритме ICP (Iterative Closest Point). Данный алгоритм используется для сведения к минимуму разницы между двумя облаками точек.

ICP часто используется для восстановления двухмерных (2D) или трёхмерных (3D) поверхностей из различных сканов, для определения местоположения роботов, детектирования моделей скелета и т.д.

2.2 Проектирование архитектуры веб-приложения

Реализованное веб-приложение придерживалось архитектуры микросервисов.

Архитектурный стиль микросервисов – это подход, при котором единое приложение строится как набор небольших сервисов, каждый из которых работает в собственном процессе и взаимодействует с остальными используя легковесные механизмы, как правило HTTP (рис. 4 Рисунок 4). Зачастую такие сервисы построены вокруг бизнес-потребностей и развертываются независимо с использованием полностью автоматизированной среды. Существует абсолютный минимум централизованного управления этими сервисами. Сами по себе эти сервисы могут быть написаны на разных языках и использовать разные технологии хранения данных.

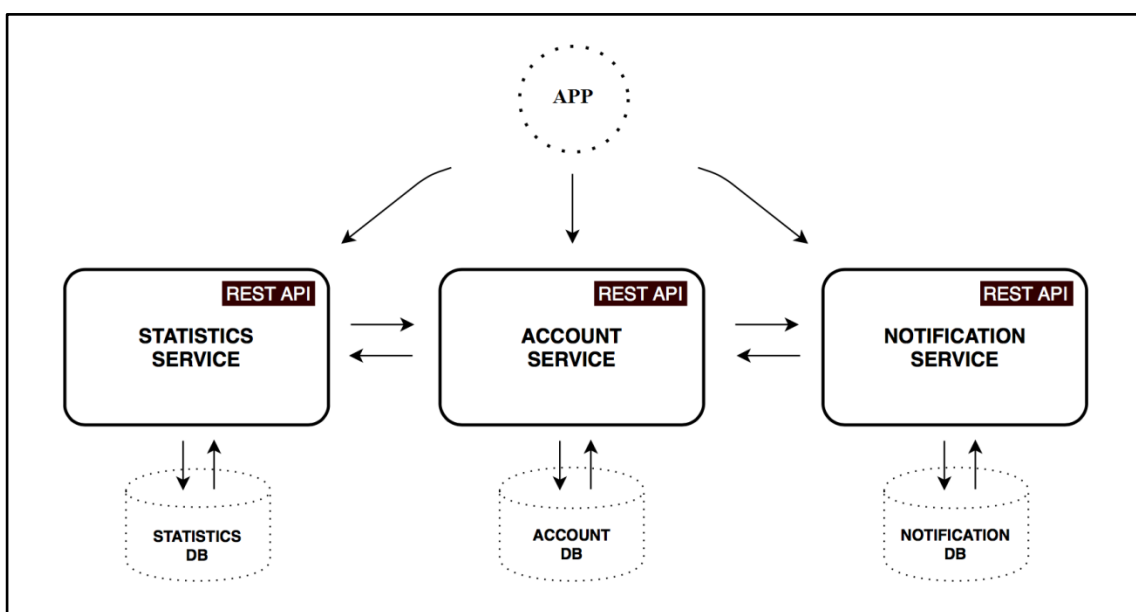


Рисунок 4 – Микросервисная архитектура

продемонстрированы с использованием диаграммы компонентов (рис. 5)

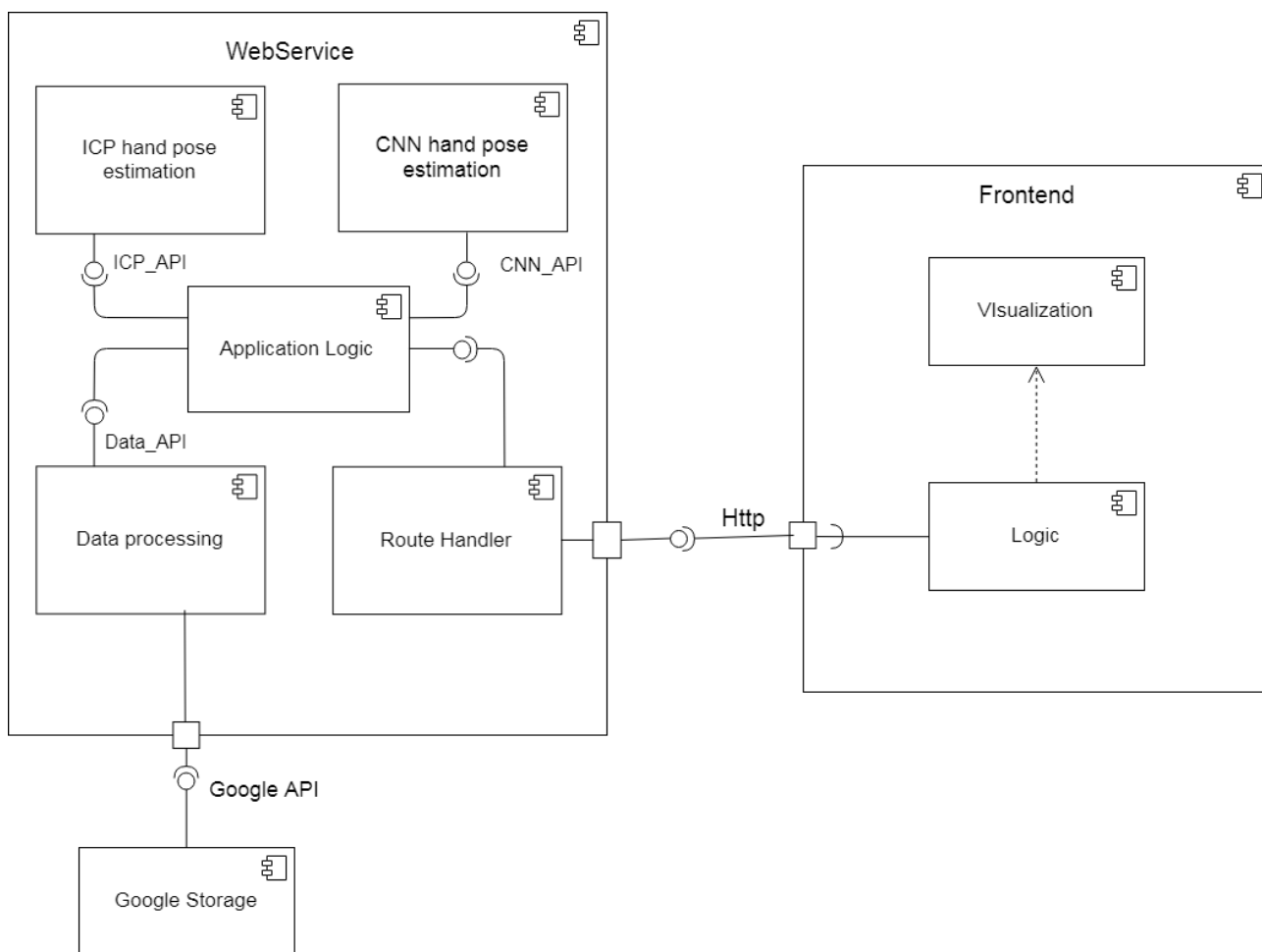


Рисунок 5 – Диаграмма компонентов

Диаграмма компонентов системы состоит из двух основных программных компонентов **WebService** и **Frontend**.

WebService предоставляет интерфейс на основе протокола **HTTP** и содержит в себе следующие компоненты:

- **Route Handler** – отвечает за обработку запросов, поступающих через **HTTP**;
- **Application Logic** – включает в себе всю бизнес-логику приложения, является связующим узлом компонента **WebService**;
- **Data Processing** – выполняет загрузку и предобработку данных, загрузка происходит с **Google Storage** с использованием предоставляемого **API**;

- ICP Hand pose estimation – выполняет оценку позы руки на основе алгоритма ICP;
- CNN Hand pose estimation – выполняет оценку позы руки на основе сверточных нейронных сетей.

Компонент Frontend отвечает за клиентскую часть системы, данный компонент взаимодействует с Webservice через предоставляемый интерфейс. Включает в себя следующие компоненты:

- Logic – содержит логику взаимодействия с клиентом;
- Visualization – выполняет визуализацию полученных результатов работы компонентов оценки позы руки.

Для представления общей конфигурации и топологии распределенной программной системы была использована диаграммы развертывания UML (рис. 6).

Диаграмма развертывания предназначена для визуализации элементов и компонентов программы, которые существуют лишь на этапе ее исполнения. При этом представляются только компоненты программы, являющиеся исполняемыми файлами или динамическими библиотеками.

Левая часть диаграммы представляет собой клиента, который с помощью web-браузера по протоколу HTTP взаимодействует с сервером.

Web-сервер представляет собой некую исполняемую среду, реализованную с помощью программной платформы Node.js и web-фреймворка Koa. В основу фреймворка положены три модуля: Business Logic, Middlewares, Route Handler. Business Logic содержит в себе всю бизнес-логику приложения. Middlewares представляет собой промежуточный уровень, осуществляющий предобработку запросов. Route Handler организует маршрутизацию клиентских запросов.

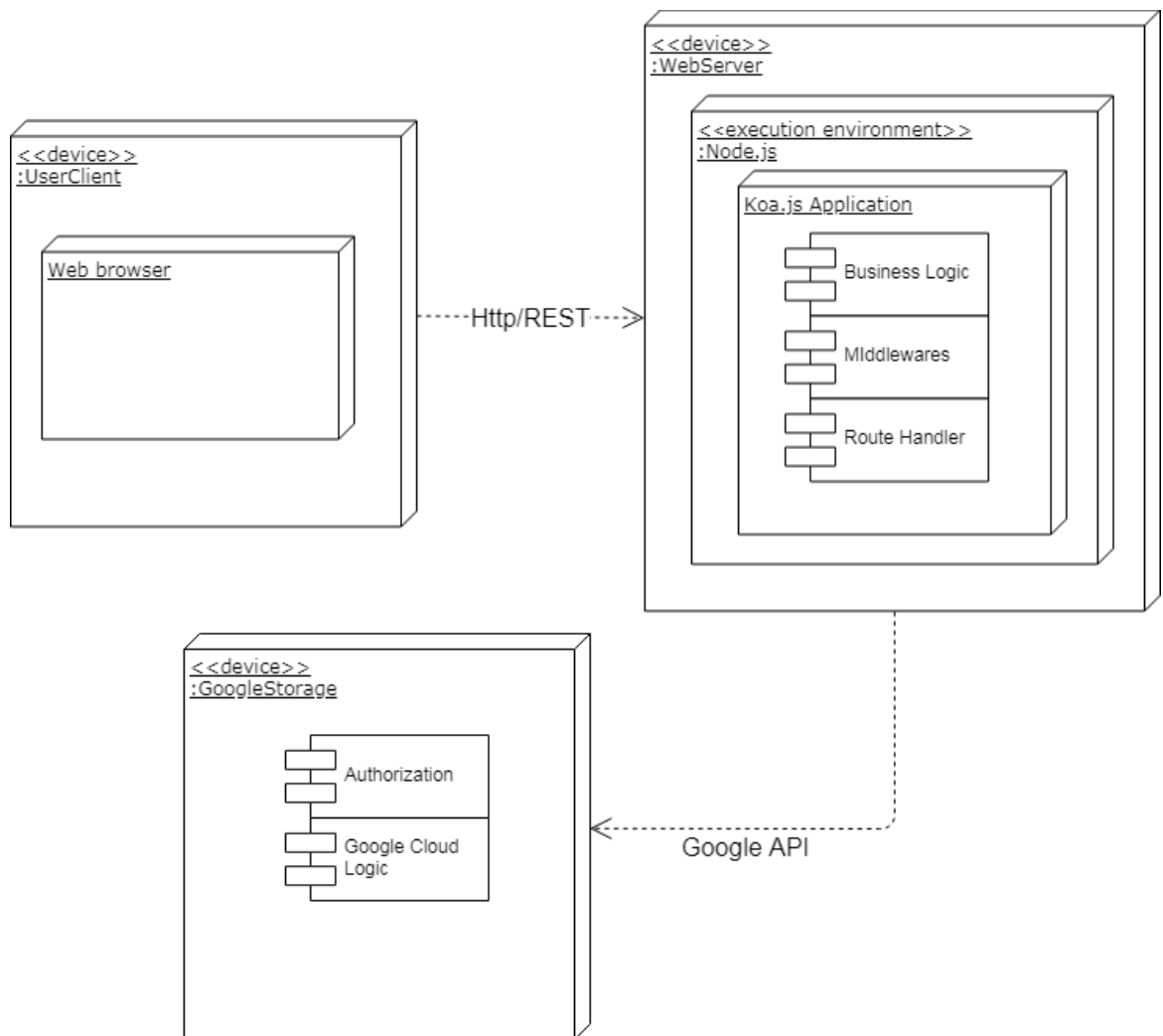


Рисунок 6 – Диаграмма развертывания

2.2.1 Преимущества и недостатки микросервисной архитектуры

К положительным сторонам можно отнести:

- четкое деление по модулям. Всегда понятно, как работает та или иная часть кода. Просто внедрять новый функционал;
- высокая доступность. Если какая-то часть монолита сломается – сломается все приложение. С микросервисами иначе: сервисы могут работать не все (не критические, вроде авторизации), но приложение при этом останется доступным;
- разнообразие стека технологий. При разработке каждого сервиса можно выбирать инструменты, которые лучше всего подойдут для конкретной бизнес-логики в этом сервисе. Например, выбрать оптимальную базу данных и удобные инструменты для работы с ней. Микросервисная архитектура также

позволяет попробовать какую-то новую технологию на отдельном сервисе, не переписывая при этом всё приложение;

- относительная простота развертывания. Каждый сервис запускается самостоятельно, что делает процесс развертывания и отладки более чистым.

Недостатки архитектуры:

- сложность разработки. Скорость разработки – высокая плата за доступность и модульность;

- сложность поддержки. Каждый микросервис нуждается в отдельном обслуживании, поэтому нужен постоянный автоматический мониторинг.

2.3 Получение облака точек и пространственная фильтрация

Процесс получения облака точек по имеющимся картам глубины не представляет сложности, если для каждой камеры известны её внешние и внутренние параметры.

Для преобразования 2D изображения необходимо последовательно для каждого пикселя изображения выполнить следующее преобразование.

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = P^{-1} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (3)$$

Где P^{-1} – обратная матрица камеры, (X, Y, Z) – координаты точки в мировой системе координат, (x, y) – координаты пикселя, z – значение пикселя (глубина). Стоит отметить, что пиксели, со значениями близкими к нулю, игнорировались, так как координаты таких точек стремятся к бесконечности и создают лишний шум в облаке точек.

На выходе получался вектор трехмерным точек, который далее подвергался пространственной фильтрации для устранения зашумленности.

Во-первых, выполнялась пороговая фильтрация, отсеивающая точки по их расстоянию от камеры (Z координате). Было выбрано значение порога 0.8, что соответствует 80 сантиметрам в реальном мире.

Во-вторых, основываясь на том факте, что интересующий объект находится непосредственно перед камерой, определялся центр масс ближайшего большого скопления точек перед камерой. По имеющемуся центру масс выполнялся второй этап пространственной фильтрации на основе ограничивающей сферы некоторого радиуса. Все точки, не попавшие в сферу, исключались из рассмотрения.

Проделанный процесс фильтрации точек помог эффективно избавиться от шума и изолировать интересующий объект.

2.4 Итеративный алгоритм ближайших точек

Итеративный алгоритм ближайших точек (ICP) – это алгоритм, используемый для минимизации разницы между двумя облаками точек. ICP часто используется для реконструкции 2D или 3D поверхностей из различных сканов, для локализации роботов и достижения оптимального планирования пути.

В данном алгоритме одно облако точек, называемое эталоном, остается фиксированным, в то время как другое – источник, преобразуется для наилучшего соответствия эталону. Алгоритм итеративно просматривает преобразования (комбинацию перемещения и поворота), необходимые для минимизации метрики ошибки, обычно расстояния от источника до облака опорных точек, такого как сумма квадратов разностей между координатами совпадающих пар. ICP является одним из широко используемых алгоритмов при выравнивании трехмерных моделей с учетом первоначального предположения о необходимости преобразования твердого тела.

Входные данные алгоритма: два облака точек, первоначальная оценка преобразования (опционально), критерий остановки итераций.

Выходные данные: искомое преобразование.

Алгоритм состоит из четырех основных шагов:

- 1) Связка точек из эталона с точками источника с использованием критерия ближайшего соседа;
- 2) Определение аффинного преобразования с использованием

оценочной функции (вычисление матрицы смещения и поворота, минимизирующей среднеквадратичную ошибку (MSE) между связанными точками);

3) Применение полученного преобразования к облаку точек и обновление результата;

4) Повторение шагов 1-3, пока не достигнут критерий останова.

2.5 Алгоритмический подход к решению задачи

Алгоритмический метод основывается на переборе возможных положений руки в связке с алгоритмов ICP, который играет роль некой оценочной функции.

Для корректного перебора жестов были определены угловые ограничения поворота суставов (табл. 1). Расчет угловых ограничений происходил на основе кинематики движения человеческой руки.

Таблица 1 – Угловые ограничения поворотов суставов в радианах

Наименование сустава	Ограничения по оси X	Ограничения по оси Y	Ограничения по оси Z
hand_skeleton_root	$[0; 2 * \pi]$	$[0; 2 * \pi]$	$[0; 2 * \pi]$
hand_skeleton_thumb_joint1	–	$[-0.5; 1.1]$	$[-0.9; 0.8]$
hand_skeleton_thumb_joint2	–	$[-0.4; 2.1]$	–
hand_skeleton_thumb_joint3	–	$[-0.8; 1.5]$	–
hand_skeleton_index_joint1	$[-0.3; 0.5]$	$[-0.3; 0.35]$	$[-1.65; 0.1]$
hand_skeleton_index_joint2	$[-0.2; 0.25]$	–	$[-1.65; 0.3]$
hand_skeleton_index_joint3			
hand_skeleton_middle_joint1	–	$[-0.2; 0.25]$	$[-1.65; 0.3]$
hand_skeleton_middle_joint2	–	–	$[-1.65; 0.3]$
hand_skeleton_middle_joint3			
hand_skeleton_ring_joint1	–	$[-0.3; 0.5]$	$[-1.65; 0.3]$
hand_skeleton_ring_joint2	–	–	$[-1.65; 0.3]$
hand_skeleton_ring_joint3			

hand_skeleton_pinky_joint1	–	[-0.5; 0.85]	[-1.65; 0.3]
hand_skeleton_pinky_joint2	–	–	[-1.65; 0.3]
hand_skeleton_pinky_joint3			

Стоит отметить, что алгоритм ICP является итеративным, то есть алгоритм постепенно приближается к наилучшему ответу. Данный факт свидетельствует о том, что алгоритм является чувствительным к начальному приближению.

Основываясь на том, что данные, полученные с сенсора, представляют собой покадрово разложенное видео, было предложено решение использовать информацию, полученную с предыдущего кадра. Таким образом, если частота съемки не будет слишком мала, то можно полагаться, что два соседних кадра будут похожи, и можно в качестве начального приближения взять результат с предыдущего кадра.

Однако исключением становится первый кадр, для которого не существует предыдущего результата. Для первого кадра было предложено вручную указать два вектора в облаке точек, для того чтобы однозначно определить направление пальцев и ладони в пространстве. Точки устанавливались в центр ладони, в направлении большого и среднего пальцев. На основе этих трех точек, строилась тройка векторов, на основании которой происходило определение направления ладони и пальцев, а также точки для выполнения начального совмещения модели и облака точек. На рисунке 7 представлен результат выполнения начального приближения модели и облака точек на основе трех точек.

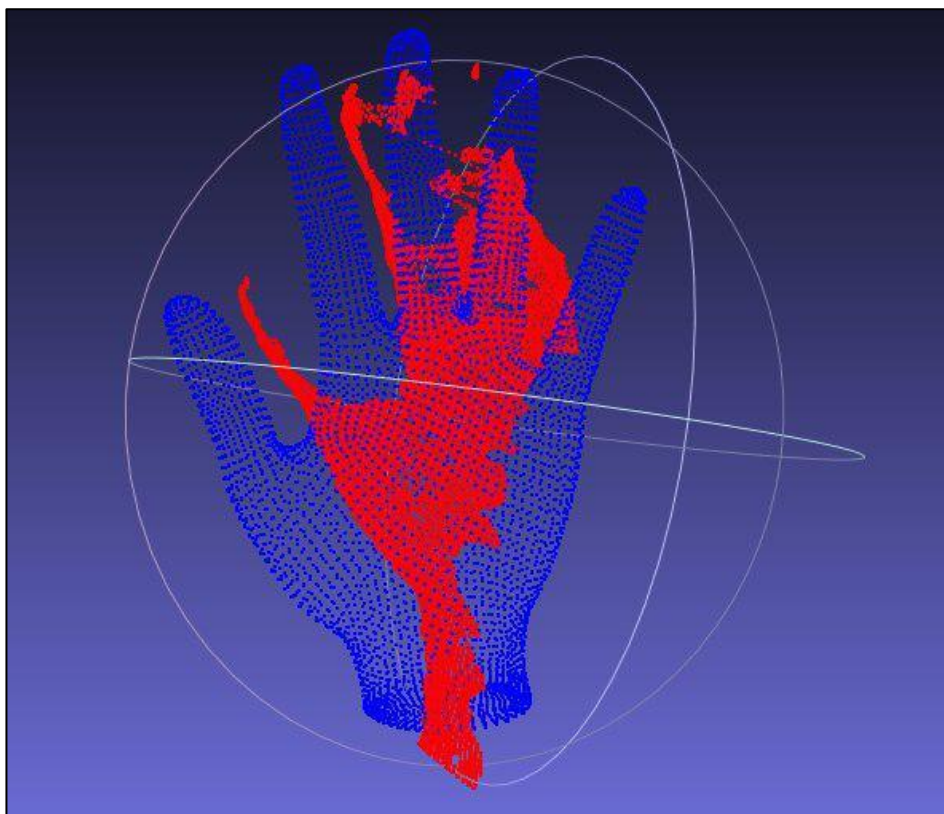


Рисунок 7 – Начальное приближение для первого кадра

Далее алгоритм можно разбить на следующие итерации:

- 1) Изменение углов поворота суставов в заданных ограничениях последовательно от большого пальца к мизинцу;
- 2) После каждого изменения угла, запуск алгоритма ICP и определение оценочной функции;
- 3) Если оценочная функция меньше, чем минимальное значение, полученное ранее, то происходит запоминание текущей трансформации и значения оценочной функции. Иначе происходит откат примененной трансформации;
- 4) Повторение шагов 1–3 пока не достигнут критерий останова.

Критерий останова срабатывает, если за полный цикл (от большого до мизинца) перебора ни разу не произошло улучшение текущего минимума значения оценочной функции.

На рисунке 8 представлен результат выполнения нескольких итераций алгоритма. Результатом работы алгоритма являются углы поворота для каждого

сустава, относительно их начального положения. Результаты сохранялись в формат JSON.

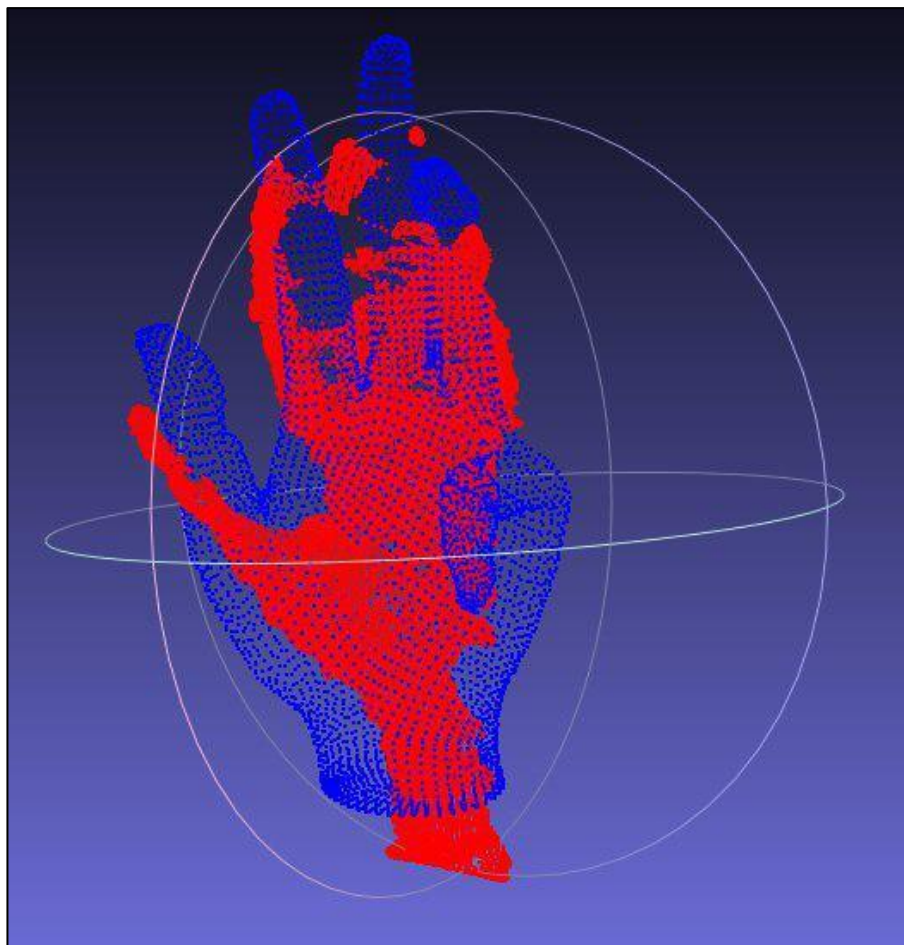


Рисунок 8 – Скелет и облако точек после выполнения нескольких итераций алгоритма

Недостатки алгоритмического подхода:

- Скорость работы, время детектирования позы для одного кадра составляло 10 – 20 секунд;
- Чувствительность к начальной инициализации;
- Необходимость в скелетной модели с известными параметрами и ограничениями движения.

2.6 Нейросетевой подход к решению задачи

В этой работе рассматривается метод 3D-регрессии, использующий CNN с проекциями в качестве входных данных. Такая архитектура сможет лучше

использовать данные с карт глубины для более точного восстановления 3D-положения суставов руки. Схема работы показана на рисунке 9.

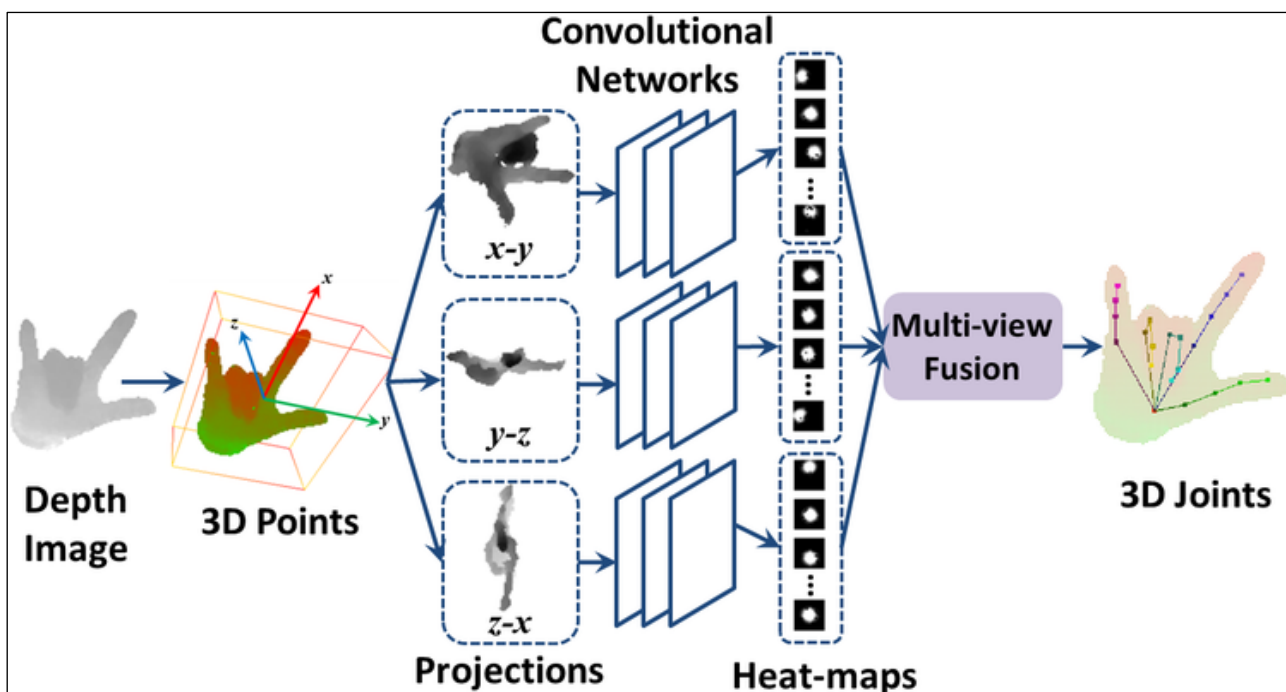


Рисунок 9 – Общая схема работы алгоритма

Облако точек, полученное из карт глубины, сначала проецируется на три ортогональных плоскости, и каждое спроецированное изображение затем подается в отдельную сверточную нейронную сеть для генерации набора тепловых карт суставов руки. Поскольку каждая тепловая карта кодирует двумерное распределение положения суставов на проецируемой плоскости, таким образом, комбинация карт в трех видах содержит распределение положения суставов в трехмерном пространстве. Объединяя тепловые карты трех видов с заранее изученными априорными позами руки, можно получить трехмерные координаты суставов и одновременно уменьшить неоднозначность оценки позы.

По сравнению с методами, использующими одну тепловую карту, предлагаемый в работе метод CNN с несколькими видами имеет следующие преимущества:

- В методах с одной тепловой картой в качестве Z координаты сустава принимается соответствующее значение из карты глубины, что может привести к большим значениям ошибки оценочной функции, даже если

прогнозируемая 2D-позиция сустава лишь незначительно отклоняется от истинной позиции, как показано на рисунке 10. Напротив, предложенный в работе вариант CNN с несколькими видами генерирует тепловые карты одновременно для переднего, бокового и верхнего видов, из которых можно более точно оценить трехмерное местоположение сустава.

– В случае неоднозначности определения местоположения сустава на тепловой карте, методы с одной тепловой картой не могут однозначно выбрать положение из нескольких горячих точек на тепловой карте, в которой только одно может соответствовать истинному положению, как показано на рисунке 11 (вид x-y). В предложенном методе с несколькими видами, тепловые карты с двух других видов могут помочь устранить такую неоднозначность.

– В отличие от других методов, в которых используется предопределенная модель руки с ограничениями положений суставов, предлагаемый в работе подход неявным образом включает ограничения позы руки, извлеченные из обучающих выборок, что позволяет применять ограничения движения руки без определения характеристик рук вручную.

Комплексные эксперименты подтверждают хорошую производительность предложенного метода на общедоступных наборах данных [10]. Кроме того, предлагаемый метод регрессии с тремя проекциями позволяет достичь относительно высокой точности в экспериментах на собственных наборах данных.

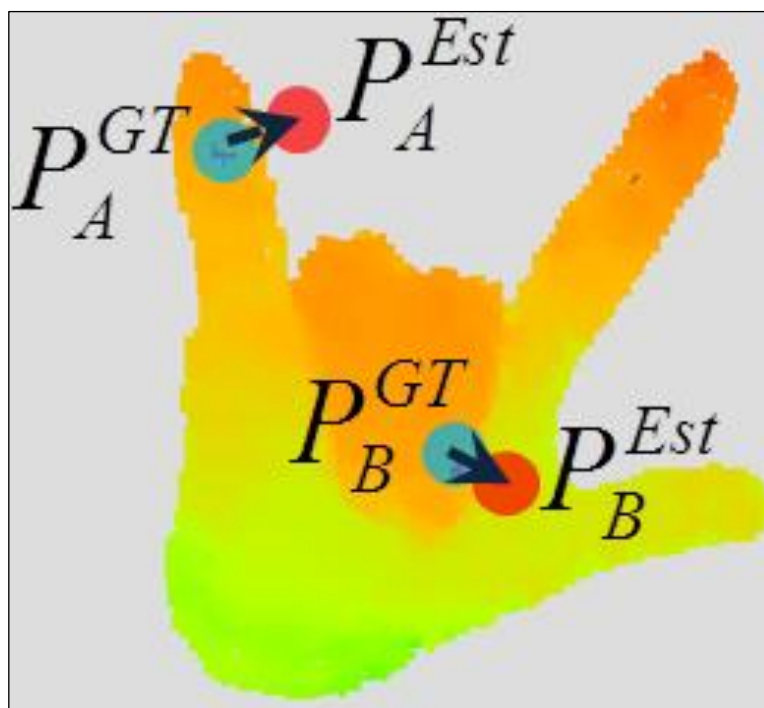


Рисунок 10 – Пример детектирования суставов (P^{EST} – полученные положения суставов, P^{GT} – истинные положения)

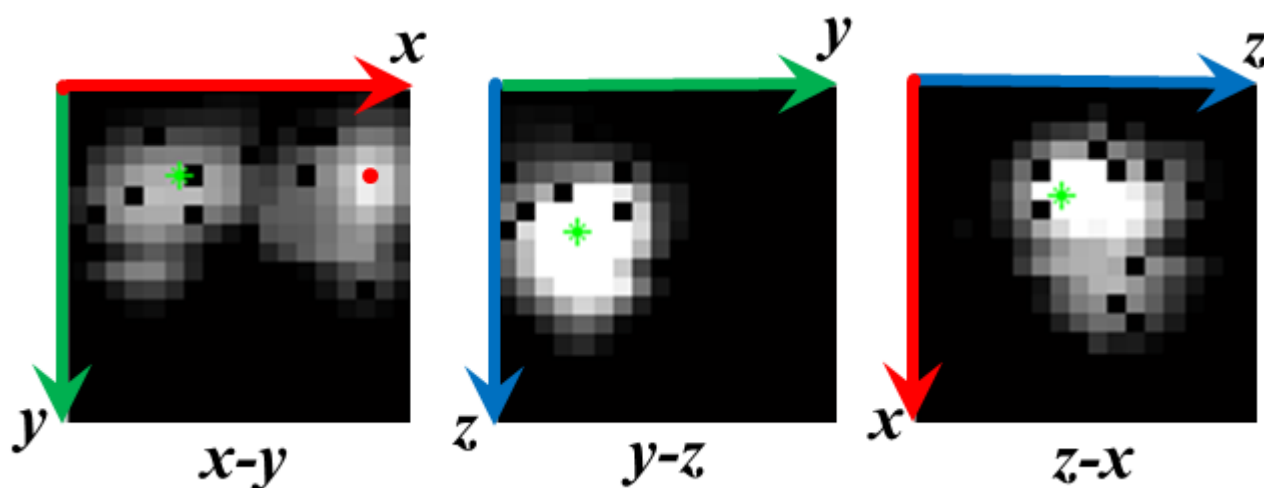


Рисунок 11 – Тепловые карты для трех проекций

2.6.1 Формализация задачи

Задача оценки позы руки может рассматриваться, как извлечение трехмерного положения сустава руки из карт глубины. В частности, входными данными для этой задачи является набор изображений глубины, содержащих только человеческую руку с некоторым жестом, а выходные данные представляют собой K векторов координат суставов в трехмерном пространстве, которые однозначно задают позу руки.

Пусть K истинных положений суставов имеют вид $\Phi = \{\phi_k\}_{k=1}^K \in \Lambda$, где Λ – пространство суставов размерности $3 \times K$. В данной работе количество суставов руки $K = 21$, это центр запястья, пять пястно-фаланговых сустава, пять проксимальных межфаланговых сустава, пять дистальных межфаланговых сустава и пять кончиков пальцев (рис. 12).

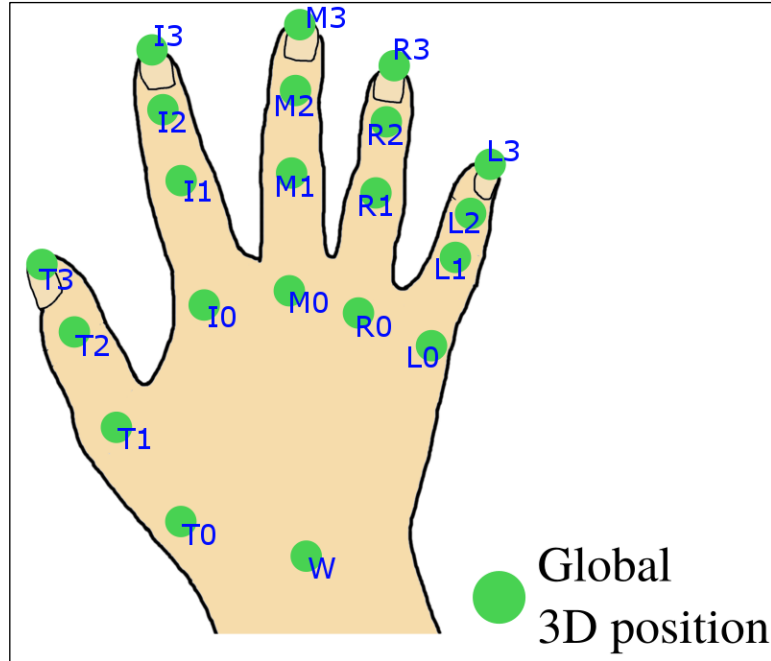


Рисунок 12 – Расположение суставов

Как говорилось в разделах выше, в предложенном методе заключение о местоположении суставов руки производится на основе проецируемых изображений в трех ортогональных плоскостях. Обозначим три проецируемых изображения I_{xy} , I_{yz} , I_{zx} , которые получаются путем проецирования трехмерного облака точек на плоскости x - y , y - z и z - x в системе координат проекции, соответственно. Таким образом, начальное изображение глубины I_D преобразуется в три проекции I_{xy} , I_{yz} и I_{zx} , которые в дальнейшем будут использоваться в качестве входных данных для определения трехмерного положения суставов руки.

Оценка местоположения сустава основывается на применении MAP, метода оптимизации на основе апостериорного максимума, на основе данных с трех проекций I_{xy} , I_{yz} и I_{zx} , которые можно рассматривать как некоторые наблюдения за 3D позой руки. Имея I_D и Φ , предполагаем, что три проекции

I_{xy} , I_{yz} и I_{zx} независимы, и обусловлены местоположениями суставов Φ . При этом предположении и предположении о равной априорной вероятности $P(\Phi)$, апостериорная вероятность местоположения суставов может быть получена, как произведение отдельных вероятностей, полученных с трех видов.

Таким образом, задача определения оптимального местоположения суставов руки формулируется следующим образом:

$$\begin{aligned}
 \Phi^* &= \arg \max_{\Phi} P(\Phi | I_{xy}, I_{yz}, I_{zx}) = \\
 &= \arg \max_{\Phi} P(I_{xy}, I_{yz}, I_{zx} | \Phi) = \\
 &= \arg \max_{\Phi} P(I_{xy} | \Phi) P(I_{yz} | \Phi) P(I_{zx} | \Phi) = \\
 &= \arg \max_{\Phi} P(\Phi | I_{xy}) P(\Phi | I_{yz}) P(\Phi | I_{zx}), \Phi \in \Omega
 \end{aligned} \tag{4}$$

Где Φ ограничивается подпространством меньшей размерности $\Omega \subseteq \Lambda$ для уменьшения неоднозначности при оценке позы.

Апостериорные вероятности $P(\phi_k | I_{xy})$, $P(\phi_k | I_{yz})$ и $P(\phi_k | I_{zx})$ могут быть вычислены по тепловым картам, генерируемые с помощью сверточных нейронных сетей.

2.6.2 Проецирование облака точек

В данном разделе подробнее рассматривается процесс получения трех проекций из облака точек.

Как показано на рисунке 9 входные карты глубины сначала преобразуются в набор трехмерных точек в мировой системе координат с использованием внутренних и внешних параметров камеры (раздел 1.3.2).

Чтобы получить проекции трех видов, облако точек проецируется на три ортогональных плоскости. Как показано на рисунке 13, ориентированный ограничивающий прямоугольник (ОВВ) генерируется путем анализа главных компонент (РСА) набора трехмерных точек. Полученный ОВВ плотно прилегает к облаку точек в локальном пространстве. Начало системы координат ОВВ задается в центре ограничивающего параллелепипеда, а её оси x , y , z соответственно выровнены с 1-ой, 2-ой и 3-ей главной компонентой. Эта система координат установлена как система координат проекции.

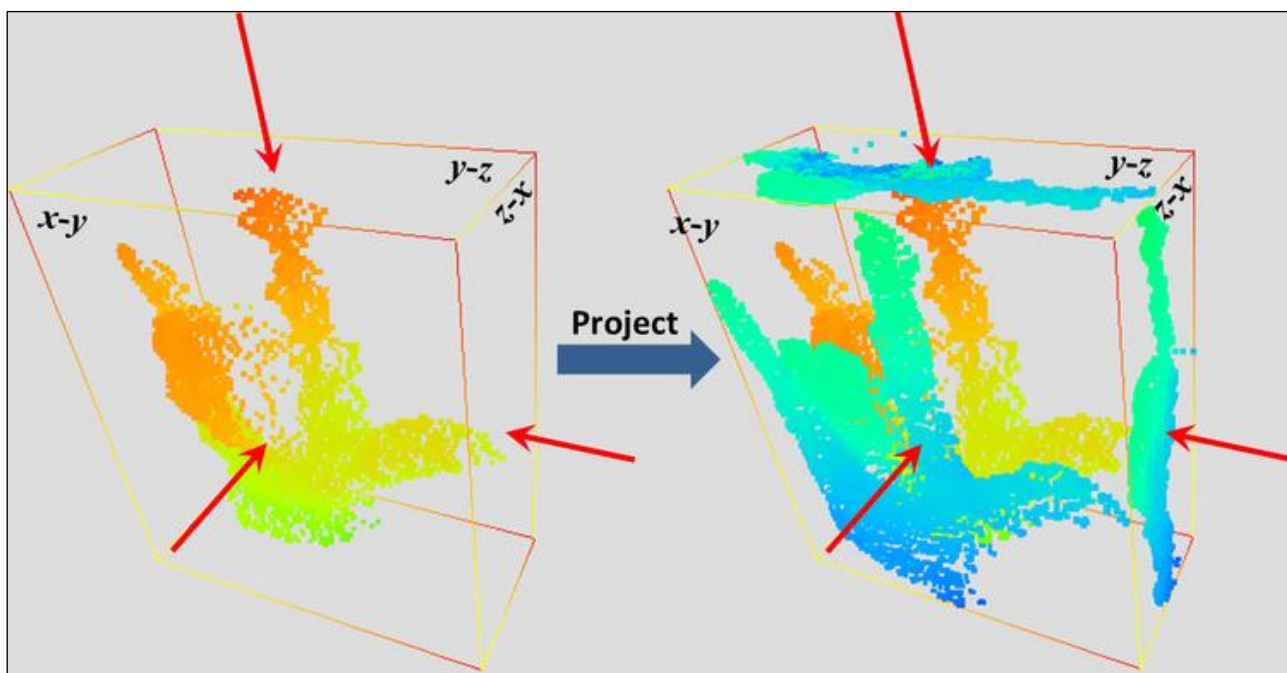


Рисунок 13 – Иллюстрация трехмерной проекции точек

Для проекции трехмерных точек на плоскость расстояния от точек до плоскости проекции нормализуются в диапазон $[0; 1]$ (для ближайших точек установлено значение 0, для самых дальних точек установлено значение 1). Затем трехмерные точки ортогонально проецируются на плоскости $x-y$, $y-z$ и $z-x$ системы координат ОВВ соответственно, как показано на рисунке 13. Соответствующие нормализованные расстояния сохраняются в виде значений пикселей проецируемых изображений. Если несколько точек проецируются на один и тот же пиксель, то в качестве значения пикселя будет сохранено наименьшее нормированное значение. Стоит отметить, что проекции на три ортогональных плоскости могут быть грубыми из-за разрешения карт глубины, данная проблема может быть решена путем применения медианного фильтра и операции размыкания к проецируемым изображениям.

2.6.3 Архитектура CNN

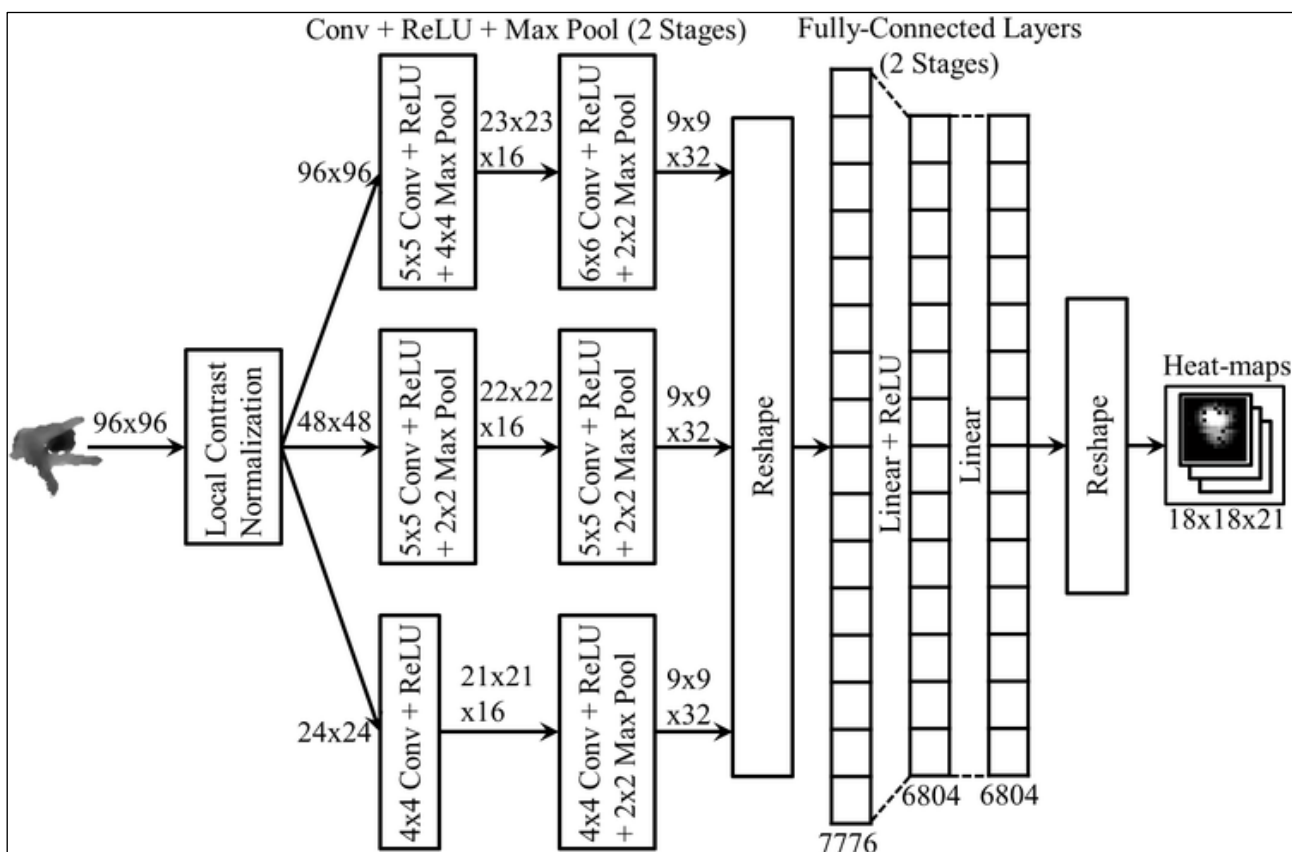


Рисунок 14 – Архитектура CNN

Поскольку имеется три изображения проекций, то для каждого вида строится собственная сверточная сеть. Все сети имеют одинаковую архитектуру и параметры. На основе изученных материалов [13], в работе используется архитектура сверточных сетей с множественным разрешением для каждого вида, как показано на рисунке 14. Размер входных изображений масштабируется до 96×96 пикселей, а затем фильтруется с помощью локальной нормализации контраста (LCN). После LCN изображение размера 96×96 подвергается понижению размерности до 48×48 и 24×24 пикселей. Все эти три изображения с разным разрешением распространяются через три блока, каждый из которых состоит из двух стадий свертки. В качестве функции активации на сверточных слоях использовалась линейная функция активации (ReLU).

Выходные карты характеристик этих трех блоков объединяются и подаются на вход полносвязной сети, содержащей один скрытый слой.

Выходной слой полносвязной сети содержит 6804 нейронов, что соответствует 21-ой тепловой карте размера 18×18 пикселей, интенсивность которых указывает на достоверность расположения сустава, в двумерных координатах, на конкретном виде.

2.6.4 Восстановление положения суставов по тепловым картам

Целью объединения результатов с нескольких видов является оценка трехмерного положения сустава по тепловым картам, полученным с использованием CNN.

Обозначим ϕ_{kx}, ϕ_{ky} и ϕ_{kz} координаты x, y, z сустава ϕ_k в системе координат ОВВ. Представленная CNN генерирует набор тепловых карт для каждого сустава. Поскольку интенсивность на тепловой карте указывает на достоверность расположения сустава в этой точке на соответствующей проекции x - y , y - z или z - x . Основываясь на тепловых картах можно получить значения следующих вероятностей $P(\phi_{kx}, \phi_{ky} | I_{yx})$, $P(\phi_{ky}, \phi_{kz} | I_{yz})$ и $P(\phi_{kz}, \phi_{kx} | I_{zx})$.

Полагая, что распределение координаты z в проекции x - y является равномерным, имеем следующее:

$$P(\phi_k | I_{xy}) = P(\phi_{kx}, \phi_{ky}, \phi_{kz} | I_{xy}) = P(\phi_{kx}, \phi_{ky} | I_{xy})P(\phi_{kz} | I_{xy}). \quad (5)$$

Отсюда следует, что вероятность $P(\phi_k | I_{xy})$ пропорциональна $P(\phi_{kx}, \phi_{ky} | I_{xy})$. Такие же выводы можно сделать для двух других проекций, получаем: $P(\phi_k | I_{yz}) \propto P(\phi_{ky}, \phi_{kz} | I_{yz})$ и $P(\phi_k | I_{zx}) \propto P(\phi_{kz}, \phi_{kx} | I_{zx})$.

Основываясь на том, что положение сустава на одной проекции не зависит от других двух. Задачу оптимизации, сформулированную в разделе 2.6.1, можно записать следующим образом:

$$\begin{aligned} \Phi^* &= \arg \max_{\Phi} P(\Phi | I_{xy})P(\Phi | I_{yz})P(\Phi | I_{zx}) = \\ &= \arg \max_{\Phi} \prod_k P(\phi_k | I_{xy})P(\phi_k | I_{yz})P(\phi_k | I_{zx}) = \\ &= \arg \max_{\Phi} \prod_k Q(\phi_{kx}, \phi_{ky}, \phi_{kz}) \end{aligned} \quad (6)$$

Где $Q(\phi_{kx}, \phi_{ky}, \phi_{kz})$ – произведение вероятностей $P(\phi_{kx}, \phi_{ky} | I_{xy})$, $P(\phi_{ky}, \phi_{kz} | I_{yz})$ и $P(\phi_{kz}, \phi_{kx} | I_{zx})$ для каждого сустава.

Полученное выше уравнение показывает, что можно получить оптимальное местоположение каждого сустава, максимизируя лишь произведение значений $Q(\phi_{kx}, \phi_{ky}, \phi_{kz})$, которые могут быть рассчитаны по интенсивностям тепловых карт трех проекций.

Для упрощения задачи вычисления произведения вероятностей, значение $Q(\phi_{kx}, \phi_{ky}, \phi_{kz})$ аппроксимируется Гауссовым распределением $N(\mu_k, \Sigma_k)$, где μ_k – вектор средних значений, а Σ_k – его ковариационная матрица. Эти параметры распределения Гаусса могут быть оценены по тестовым данным.

На основании приведенных выше предположений и выводов, задачу оптимизации можно аппроксимировать следующим образом:

$$\begin{aligned} \Phi^* &= \arg \max_{\Phi} \sum_k \log Q(\phi_{kx}, \phi_{ky}, \phi_{kz}) = \\ &= \arg \max_{\Phi} \sum_k \log N(\mu_k, \Sigma_k) = \\ &= \arg \min_{\Phi} \sum_k (\phi_k - \mu_k)^T \Sigma_k^{-1} (\phi_k - \mu_k), \end{aligned} \quad (7)$$

Где $\Phi = \sum_{m=1}^M a_m e_m + u$ принимает линейную форму. Для определения линейных ограничения конфигураций суставов Φ , выполнялся анализ главных компонент положений суставов на всей обучающей выборки. $E = [e_1, e_2, \dots, e_m]$ – вектор главных компонент, $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_m]^T$ – вектор коэффициентов главных компонент, u – эмпирическое среднее, а $M \ll 3 \times K$.

С учетом линейных ограничений наложенных на Φ , оптимальный вектор коэффициентов $\alpha^* = [\alpha_1^*, \alpha_2^*, \dots, \alpha_m^*]^T$ равен $\alpha^* = A^{-1}b$, где A – симметричная матрица размера $M \times M$, b – вектор столбец размера M .

$$A_{ij} = \sum_k e_{j,k}^T \Sigma_k^{-1} e_{i,k} \quad (8)$$

$$b_i = \sum_k (\mu_k - u_k)^T \Sigma_k^{-1} e_{i,k} \quad (9)$$

$$e_i = [e_{i,1}^T, a_{i,2}^T, \dots, a_{i,K}^T]^T; u = [u_1^T, u_2^T, \dots, u_K^T]^T; i, j = 1, 2, \dots, M.$$

Оптимальные положения суставов Φ^* восстанавливаются путем обратного проецирования оптимальных коэффициентов α^* в подпространстве Ω на исходное пространство суставов Λ .

Подводя итог, предложенный метод слияния результатов с нескольких проекций состоит из двух основных этапов. Первым шагом является оценка параметров распределения Гаусса для каждого сустава с использованием тепловых карт трех проекций. Вторым этапом является вычисление оптимальных коэффициентов α^* и восстановление оптимального местоположения суставов Φ^* . Вычисление основных компонент и вектора эмпирического среднего конфигурации суставов кисти происходит путем применения метода главных компонент (PCA) к тренировочным данным на этапе обучения.

2.6.5 Обучение CNN

Сверточная нейронная сеть, описанная в разделах выше, была реализована с использованием нейросетевой библиотеки Keras [16]. Алгоритм оптимизации, применяемый в тренировочном процессе CNN, представляет собой стохастический градиентный спуск (SGD) с функцией потерь по среднеквадратичной ошибке (MSE), поскольку задача оценки позы руки является типичной проблемой регрессии.

В качестве параметров обучения использовались следующие значения: размер обучающего пакета 64, скорость обучения 0.2, момент 0,9 и коэффициент снижения весов 0.0005. Тренировка прекращалась после 50-ти эпох, чтобы предотвратить переобучение сети. Использовалась рабочая станция с процессором Intel Core i7-4500, 64 ГБ ОЗУ и графическим процессором ASUS GeForce GTX1080Ti Turbo для обучения CNN. Обучение трех CNN можно запускать параллельно, так как они являются независимыми. Обучение одной CNN занимало около 16 часов.

Ниже представлен график зависимости точности модели от количества эпох обучения (рис. 15). В качестве конечного числа эпох было выбрано значение 50, так как после 50-ти эпох точность на тестовом наборе практически

перестает меняться, однако же, точность на тренировочном наборе растёт, что свидетельствует о переобучении.



Рисунок 15 – Зависимость точности от количества эпох

Также для определения оптимального количества эпох использовался график зависимости функции потерь от количества эпох (рис. 16). На данном графике также отчетливо видно, что после 50-ти эпох функция потерь начинает резко возрастать, это происходит из-за переобучения на тренировочном наборе.

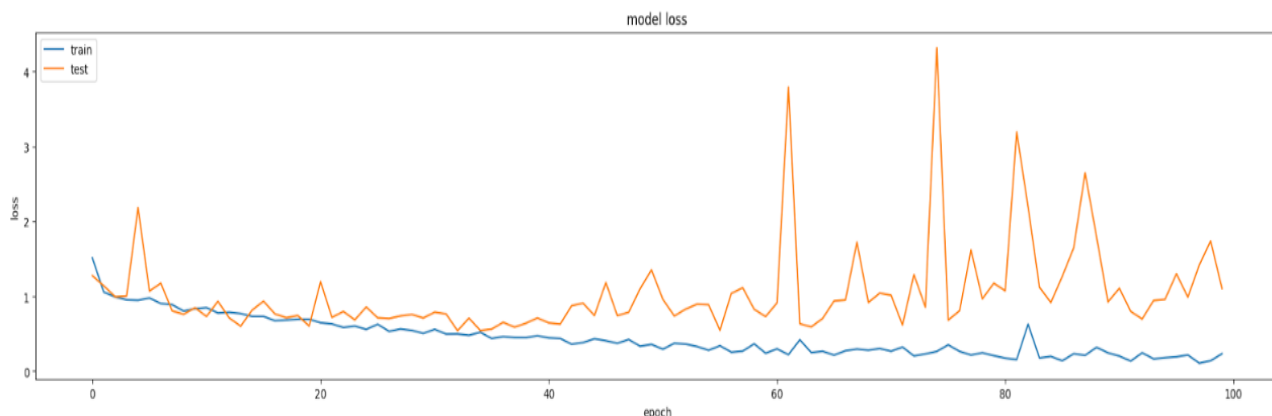


Рисунок 16 – Зависимость значения функции потерь от количества эпох

2.6.6 Наборы данных и метрики

В ходе выполнения работы производилось сравнение реализуемого подхода с современными методами для набора данных, опубликованного авторами работы [10], который является самым сложным из открытых набором данных с позами рук. Этот набор данных содержит 9 поднаборов, и каждый поднабор содержит 17 жестов, где для каждого жеста имеется по 500 тестовых примеров. При тестировании использовалось 8 поднаборов в качестве тренировочных данных для обучения CNN, а оставшийся поднабор – в качестве

тестового набора. На каждой новой итерации обучения, тестовый поднабор заменялся на один из тренировочных поднаборов.

Для оценки эффективности регрессии использовались две метрики. Первая метрика – это среднеквадратическая ошибка для каждого сустава по всем тестовым данным, данная метрика является стандартной метрикой при оценке регрессии. Второй показатель – это доля хорошо детектированных образцов во всех тестовых данных. Тестовый пример считался хорошо детектированным только тогда, когда все предполагаемые положения суставов находятся в пределах максимально допустимого расстояния от истинных положений.

2.7 Тестирование

Для тестирования были отдельно подготовлены тестовые данные. Данные подготавливались с использованием реализованного алгоритмического метода оценки позы руки, описанного в разделе 2.5.

На графике ниже представлен процент правильно детектированных поз, при изменении максимально допустимого отклонения от истинного положения суставов (рис. 17).

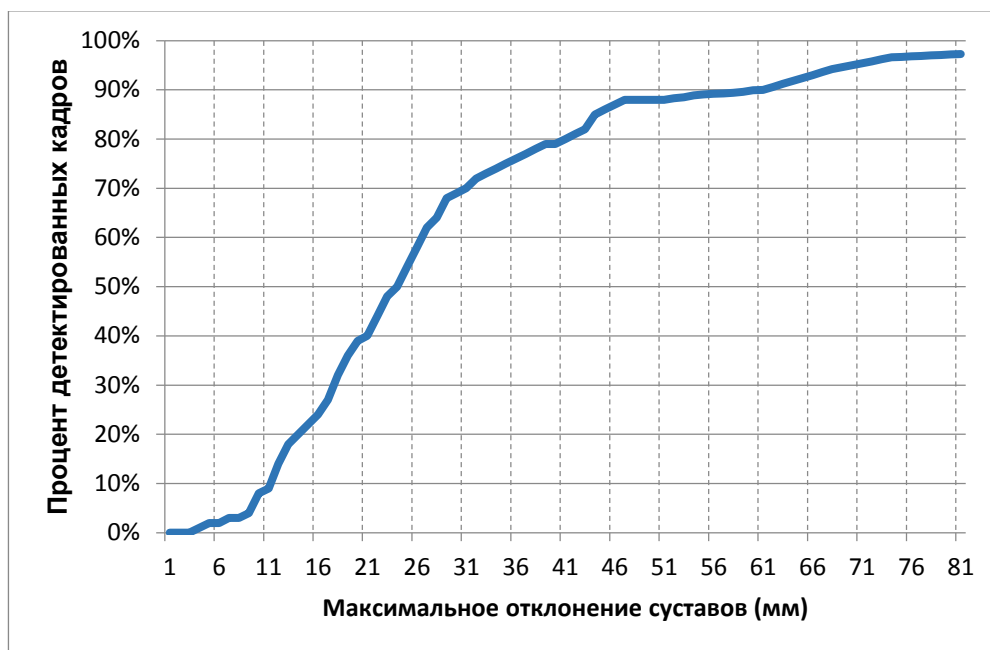


Рисунок 17 – Зависимость количества правильно детектированных кадров от значения максимального отклонения суставов

Из графика видно, что при значениях ограничения менее 20мм точность детектирования получается достаточно маленькой. Однако при ограничении примерно в 40мм точность распознавания возрастает до 80 процентов.

Чтобы проверить способность обобщения регрессионного метода, основанного на CNN, проводился эксперимент на заранее подготовленном наборе данных с шестью различными позами.

В качестве метрики использовалось среднее значение ошибки для запястья и пяти пальцев. Сравнение проводилось с двумя сторонними решениями ICP-PSO [17] и OpenPose [15].

В таблице 2 представлены результаты эксперимента. Методы ICP* и CNN* – методы, реализованные в рамках дипломной работы.

Таблица 2 – Сравнение точности

Метод	1	2	3	4	5	6	Среднее
ICP*	16.2	26.4	22.2	30.4	28.2	26.8	26.7
CNN*	30.1	19.7	24.3	19.9	21.8	20.7	22.8
ICP-PSO	10.1	24.1	13.0	12.8	11.9	20.0	15.3
OpenPose	24.2	26.1	28.0	28.4	36.2	34.1	33.9

По результатам проведенного сравнения наилучший результат показал алгоритм ICP-PSO. Однако стоит уточнить, что данный алгоритм нуждается в точной ручной модели, которая откалибрована по размеру каждого объекта съемки. В частности, для начала отслеживания эти методы используют информацию об истинном положении суставов для инициализации первого кадра.

Алгоритмический подход, предложенный в работе, показывает средние результаты, неплохо справляется с простыми жестами, но имеет высокую ошибку на сложных позах.

В такой ситуации метод на основе CNN превосходит алгоритмический метод и OpenPose, как показано в таблице 2, что указывает на то, что разработанный метод обладает хорошими способностями обобщения.

3 РЕЗУЛЬТАТЫ РАБОТЫ

В работе рассмотрена задача по разработке алгоритма автоматического восстановления позы скелета руки на основе карт глубины. В работе предложены и реализованы два подхода к решению задачи:

- 1) Алгоритмический подход с использованием алгоритма ICP;
- 2) Метод регрессии позы в 3D с использованием CNN.

Использованный подход с несколькими проекциями позволяет лучше использовать трехмерную информацию, извлекаемую из карт глубины, для точной оценки трехмерных положений суставов. Экспериментальные результаты показали, что метод показывает лучшие результаты по сравнению с алгоритмическим методом.

Далее на рисунках 18 – 21 представлены результаты работы нейросетевого алгоритма.

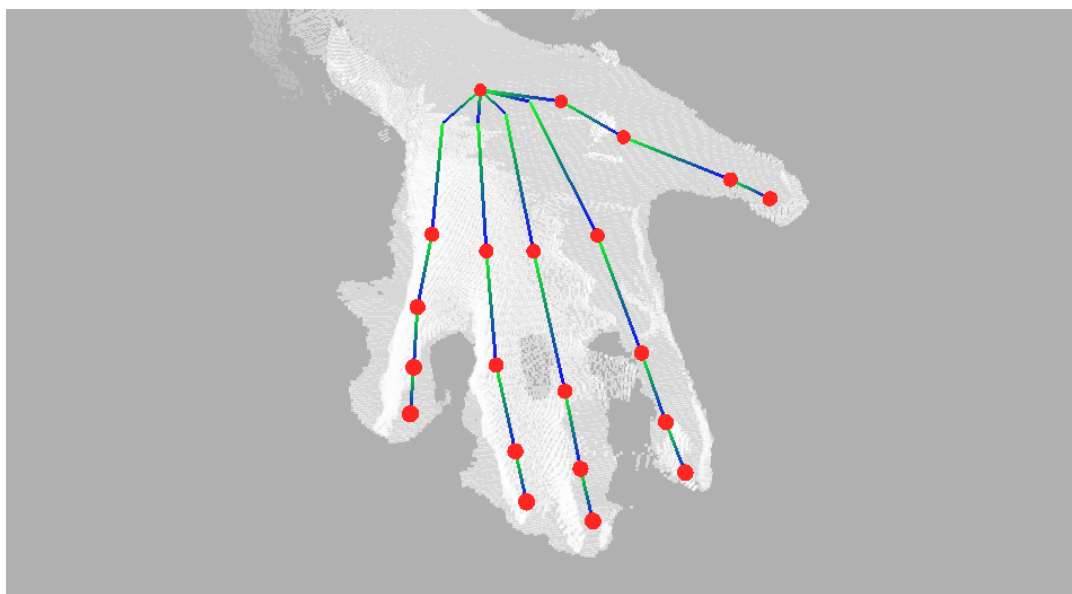


Рисунок 18 – Результаты работы нейросетевого метода

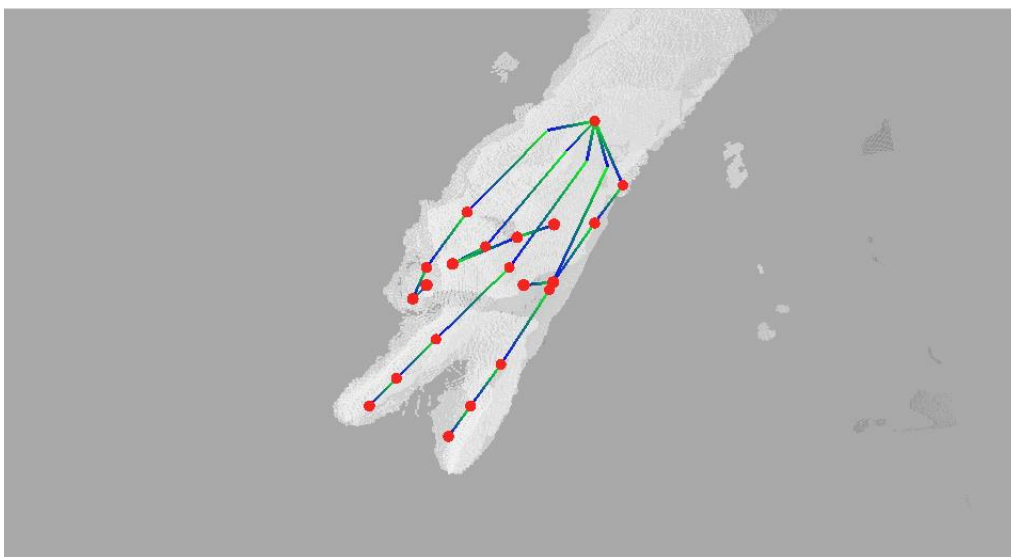


Рисунок 19 – Результаты работы нейросетевого метода

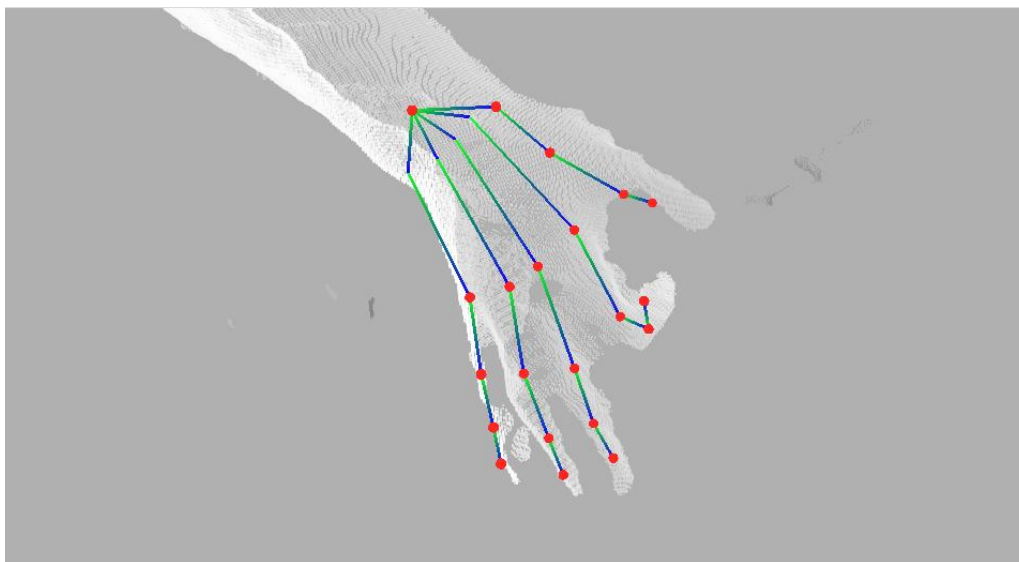


Рисунок 20 – Результаты работы нейросетевого метода

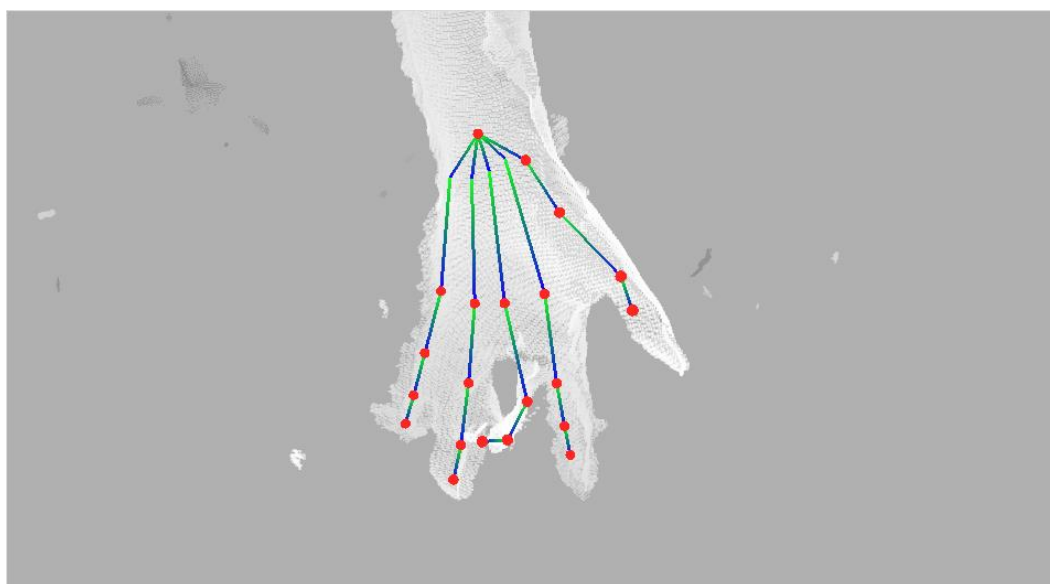


Рисунок 21 – Результаты работы нейросетевого метода

Как можно заметить из приведённых результатов, использование разработанного подхода для детектирования позы руки позволяет с высокой точностью определить местоположения суставов скелета руки.

Стоит также отметить, что результаты, полученные с использованием алгоритмического метода, на большинстве тестовых примерах не уступают по точности нейросетевому подходу. Однако время, затраченное на получение результатов для всего тестового набора с использованием нейросетевого подхода, во много раз меньше.

Также в ходе проведения сравнительного анализа предложенный метод с использованием нескольких проекций показал преимущество над методом с одной проекцией, предложенным в библиотеке OpenPose [15].

РЕСУРСОЭФФЕКТИВНОСТЬ И РЕСУРСОСБЕРЕЖЕНИЕ

Введение. В работе рассматривается задача разработки алгоритма автоматического восстановления позы скелета руки на основе карт глубины, актуальность которой обусловлена потребностью в более точных инструментах для детектирования жестов рук.

Одним из главных аспектов научного исследования, помимо перспективности является коммерческая ценность разработки, которая является необходимым условием при поиске источников финансирования.

Очевидно, что привлекать источники финансирования необходимо для проведения научного исследования и коммерциализации его результатов.

Таким образом, в разделе рассчитываются и анализируются основные показатели финансового менеджмента, ресурсоэффективности и ресурсосбережения, которые отражают состояние и перспективы проводимого научного исследования.

4.1 Потенциальные потребители результатов исследования

Произведем анализ рынка потенциальных потребителей.

Ввиду того, что разработка, которая является конечной целью научного исследования, обладает научной и практической ценностью, потенциальными потребителя являются компании, занимающиеся исследованиями или разработками в области человеко-компьютерного взаимодействия.

Главными потребителями разработки являются IT-компании, занимающиеся разработкой человеко-компьютерных интерфейсов, а также компании, работающие в области виртуальной и дополненной реальности.

4.2 Сегментирование рынка

Так как главными потребителями разработки являются IT-компании, в качестве основных критериев сегментации были выделены размер IT-компании и её территориальное расположение.

		Территориальное расположение	
		В РФ	За пределами РФ
Размер компании	Мелкие		
	Средние		
	Крупные		

Рисунок 22 – Карта сегментирования рынка услуг по использованию алгоритмов детектирования объектов:

■ анализ 2D данных ▨ анализ 2D и 3D данных

Как можно заметить из приведенной карты сегментирования рынка услуг по использованию алгоритмов детектирования объектов в видеопотоке, российский рынок является менее конкурентоспособным, а наиболее привлекательными потребителями разработки являются мелкие ИТ-компании как в России, так и за её пределами, а также средние российские ИТ-компании.

4.3 Оценка конкурентоспособности технических решений

В качестве основных конкурентных технических решений были выбраны следующие методы:

- метод детектирования на основе использование алгоритмов ICP и PSO (в таблице обозначен как «K₁»);
- метод детектирования на основе использования данных с одного изображения (в таблице обозначен как «K₂»);
- метод детектирования, предлагаемый в данной работе (в таблице обозначен как «K_ф»).

Результаты конкурентного анализа приведены в таблице 3.

Таблица 3 – Результаты конкурентного анализа

Критерии оценки	Вес	Баллы			Конкурентоспособность		
		Б ₁	Б ₂	Б _ф	К ₁	К ₂	К _ф
Технические критерии оценки ресурсоэффективности							
1. Скорость детектирования	0,3	3	4	5	0,9	1,2	1,5
2. Количество потребляемых ресурсов	0,1	4	3	3	0,4	0,3	0,3
3. Сложность реализации метода	0,05	5	4	4	0,25	0,2	0,2
4. Точность детектирования	0,3	2	4	5	0,6	1,2	1,5
5. Потенциал разработки	0,1	2	3	4	0,2	0,3	0,4
Экономические критерии оценки эффективности							
6. Доступность	0,1	4	4	3	0,4	0,4	0,3
7. Стоимость	0,05	4	3	3	0,2	0,15	0,15
Итого:		24	25	27	2,95	3,75	4,35

По результатам проведённого конкурентного анализа можно сделать вывод, что предложенный метод наиболее конкурентоспособен в сравнении с другими рассмотренными методами. В первую очередь, такая конкурентоспособность обусловлена за счёт высокой скорости работы метода, а также высокой точности детектирования.

4.4 SWOT-анализ

На основе анализа рынка и конкурентных технических решений необходимо составить матрицу SWOT-анализа, в которой показаны сильные и слабые стороны проекта, возможности и угрозы для разработки. Матрица SWOT-анализа представлена в таблице 4.

Таблица 4 – Результаты SWOT-анализа

	Сильные стороны С1.Использование современных технологий С2. Высокая точность детектирования С3.Время работы алгоритма меньше, чем у конкурентных решений	Слабые стороны Сл1. Проигрыш в точности некоторым конкурентам Сл2.Необходимость адаптироваться под различные форматы данных Сл3. Необходимость в большом количестве обучающих данных
--	--	--

Возможности В1. Дальнейшее обучение сети для повышения качества детектирования В2. Популярность и стремительное развитие направления машинного обучения В3. Рост производительности аппаратных средств В4. Создание базы обучающих примеров	В2С2: В связи с развитием направления машинного обучения и появлением новых методов и алгоритмов, возможно внедрение новых оптимизаций в данную разработку для повышения качества детектирования В3С3: С увеличением производительности аппаратных средств компьютера, замена комплектующих ведёт к увеличению скорости детектирования	В4Сл3: Создание базы данных с обучающими примерами позволит уменьшить временные затраты на подготовку данных В3Сл1: Использование более производительных аппаратных средств повысит конкурентоспособность разработки
Угрозы У1. Развитие и появление новых аналогов разрабатываемой системы У2. Прекращение поддержки и дальнейшего развития разработки У3. Нехватка средств для улучшения используемого оборудования	У2С2С3: Повышение конкурентоспособности разработки с акцентированием внимания на достоинствах У3С3: Снижение количества потребляемых ресурсов за счёт использования современных технологий	У1Сл1: Постоянный анализ появляющихся аналогов, изучение их особенностей и своевременное улучшение своей разработки

На втором этапе проведения SWOT-анализа проводится составление интерактивных матриц проекта, в которых производится анализ соответствия параметров SWOT каждого с каждым. Соотношения параметров представлены в таблицах 5-8.

Таблица 5 – Сильные стороны и возможности

	Сил1	Сил2	Сил3
В1	-	-	-
В2	-	+	-
В3	-	-	+

Таблица 6 – Слабые стороны и возможности

	Слаб1	Слаб2	Слаб3
B1	-	-	+
B2	-	-	-
B3	+	-	-

Таблица 7 – Сильные стороны и угрозы

	Сил1	Сил2	Сил3
У1	-	-	-
У2	+	+	-
У3	-	-	+

Таблица 8 – Слабые стороны и угрозы

	Слаб1	Слаб2	Слаб3
У1	+	-	-
У2	+	-	-

Таким образом, можно сделать вывод, что проект необходимо развивать в направлении наибольшей универсальности и гибкости, с возможностью расширения с помощью дообучения модели на новых данных обучения. При этом необходимо отслеживать и, по возможности, применять новейшие технологии и улучшения в области нейросетевых алгоритмов.

4.5 Оценка готовности проекта к коммерциализации

Для оценки готовности научной разработки к коммерциализации была заполнена таблица 9, содержащая показатели о степени проработанности проекта с позиции коммерциализации и компетенциям разработчика научного проекта.

Таблица 9 – Бланк оценки степени готовности научного проекта к коммерциализации

№ п/п	Наименование	Степень проработанности научного проекта	Уровень имеющихся знаний у разработчика
1.	Определен имеющийся научно-технический задел	4	4
2.	Определены перспективные направления коммерциализации научно-технического задела	4	4

3.	Определены отрасли и технологии (товары, услуги) для предложения на рынке	4	4
4.	Определена товарная форма научно-технического задела для представления на рынок	3	3
5.	Определены авторы и осуществлена охрана их прав	3	2
6.	Проведена оценка стоимости интеллектуальной собственности	2	1
7.	Проведены маркетинговые исследования рынков сбыта	2	2
8.	Разработан бизнес-план коммерциализации научной разработки	1	1
9.	Определены пути продвижения научной разработки на рынок	2	2
10.	Разработана стратегия (форма) реализации научной разработки	2	1
11.	Проработаны вопросы международного сотрудничества и выхода на зарубежный рынок	2	1
12.	Проработаны вопросы использования услуг инфраструктуры поддержки, получения льгот	2	1
13.	Проработаны вопросы финансирования коммерциализации научной разработки	2	1
14.	Имеется команда для коммерциализации научной разработки	2	1
15.	Проработан механизм реализации научного проекта	2	3
	ИТОГО БАЛЛОВ	37	31

Оценка готовности научного проекта к коммерциализации (или уровень имеющихся знаний у разработчика) определяется по формуле:

$$B_{\text{сум}} = \sum B_i ,$$

где $B_{\text{сум}}$ – суммарное количество баллов по каждому направлению;

B_i – балл по i -му показателю.

После проведения анализа, можно сделать вывод, что перспективность проведения коммерциализации средняя. Можно сделать вывод, что разработка

находится на стадии реализации и вопросы, связанные с коммерциализацией, плохо проработаны.

4.5.1 Методы коммерциализации результатов научно-технического исследования

Для успешной коммерциализации научно-технических разработок необходимо правильно определить метод коммерциализации. Данный метод определяет успешность продвижения продукта на рынке, коммерческий эффект которого может быть использован для дальнейшего продвижения и развития проекта.

Организация совместных предприятий, работающих по схеме «российское производство – зарубежное распространение» – является наиболее подходящим методом коммерциализации данной разработки. Так как область применения разработки в основном сосредоточена на зарубежном рынке, то открытие такого предприятия может принести большой доход.

Также возможен вариант коммерциализации на основе передачи «ноу-хау», так как результатом разработки является программный продукт, то возможна передача программного продукта третьему лицу на основе договора.

4.6 Инициация проекта

Инициация проекта состоит из процессов, которые выполняются для нового проекта или новой стадии проекта. Для этого определяются начальные цели, содержание, фиксируются ресурсы. Также определяются внутренние и внешние заинтересованные стороны проекта.

4.6.1 Цели и результат проекта

В данном разделе необходимо привести информацию о заинтересованных сторонах проекта, иерархии целей проекта и критериях достижения целей.

Под заинтересованными сторонами проекта понимаются лица или организации, которые активно участвуют в проекте или интересы которых могут быть затронуты как положительно, так и отрицательно в ходе

исполнения или в результате завершения проекта. Это могут быть заказчики, спонсоры, общественность и т.п. Информация по заинтересованным сторонам проекта представлена в таблице 10.

Таблица 10 – Заинтересованные стороны проекта

Заинтересованные стороны проекта	Ожидания заинтересованных сторон
ООО «Рубиус Групп»	Внедрение данной разработки в комплексную систему по подготовке данных для нейронных сетей
Компании, занимающиеся разработкой устройств виртуальной реальности	Детектирование рук позволит более качественно реализовать взаимодействие человека и системы в рамках виртуальной реальности
Компании, разрабатывающие устройства для работы за компьютером людей с ограниченными возможностями	Разработка позволит улучшить качество взаимодействия устройств с людьми из данной категории
Компании, работающие с нейронными сетями	Качественное детектирование рук позволит быстрее подготавливать обучающие данные

Исходя из ожиданий заинтересованных сторон, была сформирована цель и ожидаемые результаты проведения исследования (таблица 11).

Таблица 11 – Цели и результат проекта

Цели проекта:	Разработка алгоритма автоматического восстановления позы скелета руки на основе карт глубины
Ожидаемые результаты проекта:	Алгоритм, позволяющий детектировать положение руки в пространстве, со скоростью не менее 20 кадров в секунду и погрешностью не более 5мм.
Критерии приемки результата проекта:	Основным критерием является время обработки одного кадра, то есть время, требуемое для детектирования позы на кадре. Также стоит отметить точность детектирования, которая выражается отклонением полученной позы от истинного

	положения руки.
Требования к результату проекта:	Требование:
	Разработанный алгоритм должны позволять детектировать жесты в видеопотоке со скоростью не менее 20 кадров в секунду.
	Значение отклонения детектированного положения руки от истинного положения не должно превышать 5 мм для каждого сустава.

На данном этапе были определены заинтересованные стороны проекта, были определены цели и ожидаемые результаты проекта, а также критерии приемки результатов проекта заинтересованной стороной.

4.6.2 Организационная структура проекта.

При проведении сложных научно-исследовательских работ необходимо организовать группу способную разработать сам проект, оценить возможные экономические и социальные риски проекта и выполнить требования потенциальных потребителей.

Для выполнения вышенаписанных требований сформирована следующая группа (таблица 12).

Таблица 12 – Рабочая группа проекта

№ п/п	ФИО, основное место работы, должность	Роль в проекте	Функции	Трудо-затраты, час.
1.	Спицын В.Г., профессор отделения ИТ ТПУ	Руководитель проекта	Составление ТЗ, консультации, проверка полученных результатов	124
2.	Иовлев Ю.Д., магистрант ТПУ	Исполнитель по проекту	Анализ предметной области, разработка архитектуры, проектирование, реализация, тестирование, подготовка документации	1218
				1342

4.6.3 Ограничения и допущения проекта.

Ограничения и допущения сведены в таблице 13.

Таблица 13 – Ограничения проекта

Фактор	Ограничения/ допущения
3.1. Бюджет проекта	389379
3.1.1. Источник финансирования	НИ ТПУ
3.2. Сроки проекта:	03.09.2018 – 02.06.2019
3.2.1. Дата утверждения плана управления проектом	03.09.2018
3.2.2. Дата завершения проекта	02.06.2019
3.3. Прочие ограничения и допущения*	-

4.7 График проведения и бюджет НТИ

Перечень этапов и работ в рамках проведения научного исследования представлен в таблице 14 в виде календарного плана проекта:

Таблица 14 – Календарный план

№	Название	Продолжительность, дни	Дата начала работ	Дата окончания работ	Состав участников
1	Составление и утверждение задач	14	03.09.18	18.09.18	Руководитель
2	Составление технической документации	8	19.09.18	27.09.18	Руководитель, инженер
3	Анализ литературы в исследуемой области	36	28.09.18	09.11.18	Инженер
4	Обзор существующих решений в исследуемой области	27	10.11.19	11.12.18	Инженер
5	Календарное планирование работ по теме	10	12.12.18	22.12.18	Руководитель, инженер
6	Разработка алгоритма, решающего поставленную задачу	27	23.12.18	01.02.19	Инженер
7	Подготовка результатов для выполнения анализа	11	02.02.19	14.02.19	Инженер

8	Оценка результатов и дальнейшая работа над улучшением алгоритма	15/37	15.02.19	01.04.19	Руководитель, инженер
9	Сравнение эффективности алгоритма с аналогами	11	02.04.19	14.04.19	Инженер
10	Анализ и оформление полученных результатов	15	15.04.19	05.05.19	Руководитель, инженер
11	Оформление документации	21	6.05.19	02.06.19	Инженер

Таблица 15 – Календарный план-график работ

№	Вид работ	$T_{k,i}$	Продолжительность выполнения работ																																							
			Сент			Окт			Ноябрь			Дек			Янв.			Февр			Март			Апр			Май			Июнь												
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3													
1	Составление и утверждение технического задания	15																																								
2	Составление технической документации	8																																								
3	Анализ литературы в исследуемой области	42																																								
4	Обзор существующих решений	31																																								
5	Календарное планирование работ по теме	10										 																														
6	Разработка алгоритма, решающего поставленную задачу	40																																								
7	Подготовка результатов для выполнения анализа	12																																								

Продолжение таблицы 15

№	Вид работ	$T_{k,i}$	Продолжительность выполнения работ																													
			Сент			Окт			Ноябрь			Дек			Янв			Февр			Март			Апр			Май			Июнь		
			1	2	3	1	1	2	3	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3			
8	Оценка результатов и дальнейшая работа над улучшением алгоритма	44																														
9	Сравнение эффективности алгоритма с аналогами	12																														
10	Анализ и оформление полученных результатов	20																														
11	Оформление документации	21																														

Примечание к таблице 15:

 – Руководитель

 – Инженер

4.7.1 Бюджет НТИ

В состав затрат на создание проекта включается величина всех расходов, необходимых для реализации комплекса мероприятий, составляющих содержание разработки. Расчёт стоимости выполнения работы производится по следующим статьям затрат:

- заработная плата;
- накладные расходы;
- прочие прямые расходы;
- специальное оборудование для научных работ (амортизация);
- отчисления на социальные нужды.

Амортизация оборудования

В данной статье рассчитывается амортизация используемого оборудования за время выполнения научно-исследовательской работы:

$$C_{am} = \frac{S_n}{T} * T_f,$$

где S_n – балансовая стоимость единицы оборудования;

T – срок полезного использования оборудования (в днях);

T_f – фактическое время работы оборудования (в днях);

Таблица 16 – Затраты на амортизацию оборудования

Наименование оборудования	Балансовая стоимость единицы оборудования, руб.	Срок полезного использования оборудования, в днях	Фактическое время работы, в днях	Затраты, руб.
Персональный компьютер	25 000,00	1095	203	11865

Затраты на амортизационные отчисления составляют:

$$C_{am} = \frac{64000}{365*3} * 203 = 11865 \text{ руб.}$$

Заработная плата

Статья расходов включает заработную плату научного руководителя и студента, а также премии, входящие в фонд заработной платы. Расчет основной заработной платы выполняется на основе трудоёмкости выполнения каждого этапа и величины месячного оклада исполнителя.

Заработная плата рассчитывается по формуле:

$$C_{зп} = З_{осн} + З_{доп},$$

где $З_{осн}$ – основная заработная плата; $З_{доп}$ – дополнительная заработная плата.

Основная заработная плата рассчитывается по формуле:

$$З_{осн} = З_{дн} * T_{раб},$$

где $З_{дн}$ – среднедневная заработная плата работника, руб;

$T_{раб}$ – продолжительность работ, выполняемых работником, раб.дн.

Среднедневная заработная плата рассчитывается по формуле:

$$З_{дн} = \frac{З_{м} * M}{F_{д}},$$

где $З_{м}$ – месячный должностной оклад работника, руб;

$F_{д}$ – действительный годовой фонд рабочего времени, раб. дн;

M – количество месяцев работы без отпуска в течение года: при отпуске в 28 раб. дня $M = 11,2$ месяца, 5-дневная неделя; при отпуске в 56 раб. дней $M = 10,4$ месяца, 6-дневная неделя.

Среднедневная заработная плата руководителя:

$$З_{дн} = \frac{91853 * 10,4}{251} = 3806 \text{ руб}$$

Среднедневная заработная плата инженера (на предприятии):

$$З_{дн} = \frac{16462 * 10,4}{251} = 682 \text{ руб}$$

Таблица 17 – Баланс рабочего времени

Показатели рабочего времени	Руководитель	Инженер
Календарное число дней	365	365
Количество нерабочих дней:		
- выходные дни;	52	52
- праздничные дни.	14	14
Потери рабочего времени:		
- отпуск;	48	30
- невыходы по болезни.		18
Действительный годовой фонд рабочего времени	251	251

Расчёт месячного должностного оклада:

$$З_{\text{м}} = З_{\text{б}} * (1 + K_{\text{пр}} + K_{\text{д}}) * K_{\text{р}},$$

где $З_{\text{б}}$ – базовый оклад, руб;

$K_{\text{пр}}$ – премиальный коэффициент, принят за 30% от $З_{\text{б}}$ для преподавателя и 10% от $З_{\text{б}}$ инженера на предприятии;

$K_{\text{д}}$ – коэффициент доплат и надбавок, принят за 20% от $З_{\text{б}}$ для преподавателя;

$K_{\text{р}}$ – районный коэффициент, 1.3 для Томска.

Расчёт месячного должностного оклада руководителя:

$$З_{\text{м}} = 47104 * (1 + 0,3 + 0,2) * 1,3 = 91853 \text{ руб.}$$

Расчёт месячного должностного оклада студента на предприятии:

$$З_{\text{м}} = 12663 * 1,3 = 16462 \text{ руб.}$$

Таблица 18 – Расчёт основной заработной платы

Исполнители	$З_{\text{б}}$	$K_{\text{пр}}$	$K_{\text{д}}$	$K_{\text{р}}$	$З_{\text{м}}$, руб	$З_{\text{дн}}$, руб	$T_{\text{р}}$, раб.дн.	$З_{\text{осн}}$, руб
Руководитель	47104	0,3	0,2	1,3	91853	3806	20,7	78784
Инженер	12663	0	0	1,3	16462	682	203	138446
Итого $З_{\text{осн}} = 217230$ руб								

Дополнительная заработная плата рассчитывается исходя из 10-15% от основной заработной платы, работников.

$$З_{\text{доп}} = З_{\text{осн}} * K_{\text{доп}},$$

где $З_{\text{осн}}$ – основная заработная плата;

$K_{\text{доп}}$ – коэффициент дополнительной заработной платы принимается равным 0.1.

Дополнительная заработная плата руководителя:

$$З_{\text{доп}} = 78784 * 0.1 = 7878 \text{ руб}$$

Дополнительная заработная плата инженера на предприятии:

$$З_{\text{доп}} = 138446 * 0.1 = 13845 \text{ руб}$$

Таблица 19 – Заработная плата исполнителей

Заработная плата, руб	Руководитель	Инженер
Основная	78784	138446
Дополнительная	7878	13845
Итого	86662	152291

Отчисления на социальные нужды

В данной статье расходов отражаются обязательные отчисления по установленным законодательством Российской Федерации нормам:

- государственное социальное страхование (ФСС);
- пенсионный фонд (ПФ);
- медицинское страхование (ФФОМС).

$$С_{\text{внеб}} = K_{\text{внеб}} * (З_{\text{осн}} + З_{\text{доп}}),$$

где $K_{\text{внеб}}$ – коэффициент отчислений на уплату во внебюджетные фонды, принимается равным 27,1% (в 2019 г.).

Отчисления для руководителя: $86662 \times 0,271 = 23485 \text{ руб.}$

Отчисления для инженера: $152291 \times 0,271 = 41271 \text{ руб.}$

Накладные расходы

Затраты, которые не попали в предыдущие статьи расходов, учитываются в качестве накладных расходов.

$$C_{\text{накл}} = K_{\text{накл}} * (З_{\text{осн}} + З_{\text{доп}}),$$

где $K_{\text{накл}}$ – коэффициент накладных расходов, принимается равным 0,3.

Общие накладные расходы:

$$C_{\text{накл}} = 0,3 * (152291 + 86662) = 71686 \text{ руб.}$$

Прочие прямые расходы

В качестве прочих расходов, в работе рассчитываются затраты на электроэнергию, потребляемую компьютером.

Затраты на электроэнергию можно рассчитать следующим образом:

$$\mathcal{E}_{\text{об}} = P_{\text{об}} \cdot \mathcal{C}_\mathcal{E} \cdot t_{\text{об}} = 0.3 * 1218 * 5.8 = 2119 \text{ руб.}$$

Бюджет разработки

Определим итоговые затраты на научно-технический проект по разработке алгоритма автоматического восстановления позы скелета руки на основе карт глубины.

Таблица 20 – Бюджет затрат НТИ

№	Статья расходов	Сумма, руб	% от общей суммы
1	Амортизация оборудования	11865	3,1
2	Заработная плата	238953	61,4
3	Отчисления на социальные нужды	64756	16,6
4	Накладные расходы	71686	18,4
5	Прочие прямые расходы	2119	0,5
	Итого	389379	100

4.8 Реестр рисков проекта

Идентифицированные риски проекта включают в себя возможные неопределенные события, которые могут возникнуть в проекте и вызвать последствия, которые повлекут за собой нежелательные эффекты. Собранная информация отражена в таблице 21.

Таблица 21 – Реестр рисков

№	Риск	Потенциальное воздействие	Вероятность наступления (1-5)	Влияние риска (1-5)	Уровень риска	Способы смягчения риска	Условия наступления
1	Снижение конкурентоспособности	Прекращение разработки проекта	4	4	Высокий	Постоянный мониторинг научных статей и исследований в предметной области	Появление разработок, имеющих лучшие показатели производительности и точности
2	Прекращение поддержки и дальнейшего развития разработки	Прекращение разработки проекта	3	5	Средний	Поиск заказчиков, готовых инвестировать или покупать данную разработку	Отсутствие заказчиков на данную разработку
3	Потеря актуальности данной разработки	Прекращение разработки проекта	3	4	Низкий	Вести дальнейшую разработку в сторону универсальности, что позволит разработке выполнять более обширный спектр задач	Смена вектора развития области дополненной и виртуальной реальности

4.9 Оценка ресурсной, финансовой и экономической эффективности НТИ

Интегральный показатель ресурсной эффективности вариантов исполнения объекта исследования можно определить следующим образом:

$$I_{pi} = \sum a_i \cdot b_i ,$$

где a_i – весовой коэффициент i -го варианта исполнения;

b_i – балльная оценка i -го варианта исполнения разработки, устанавливаемого экспертным путем по выбранной шкале оценивания.

Таблица 22 – Сравнительная оценка характеристик вариантов исполнения проекта

Критерии оценки	Вес	Баллы		
		Аналог 1	Аналог 2	Текущий проект
1. Скорость детектирования	0,3	3	4	5
2. Количество потребляемых ресурсов	0,1	4	3	3
3. Сложность реализации метода	0,15	5	4	4
4. Точность детектирования	0,3	2	4	5
5. Потенциал разработки	0,15	2	3	4
Итого:	1	2,95	3,75	4,35

Сравнение интегрального показателя эффективности, вариантов исполнения разработки позволяет определить сравнительную эффективность проекта таблица 23 и выбрать наиболее целесообразный вариант из предложенных.

Сравнительная эффективность проекта ($\mathcal{E}_{\text{ср}}$):

$$\mathcal{E}_{\text{ср}} = \frac{I_{\text{исп1}}}{I_{\text{исп2}}}$$

Таблица 23 – Сравнительная эффективность разработки

№ п/п	Показатели	Аналог 1	Аналог 2	Текущий проект
1	Интегральный финансовый показатель разработки	1	1	1
2	Интегральный показатель ресурсоэффективности разработки	2,95	3,75	4,35
3	Интегральный показатель эффективности	2,95	3,75	4,35
4	Сравнительная Эффективность вариантов исполнения	$A_{н1}/A_{н2}=0,79$ $A_{н1}/ТП=0,68$	$A_{н2}/A_{н1}=1,27$ $A_{н2}/ТП=0,86$	$ТП/A_{н1}=1,47$ $ТП/A_{н2}=1,16$

Из полученных и приведённых в таблице 23 результатов, можно сделать вывод, что решение, предлагаемое в магистерской диссертации, превосходит рассмотренные аналоги. При этом преимущество над вторым аналогом незначительное.

Заключение. В ходе оценки перспективности и альтернатив проведения научного исследования с позиции ресурсоэффективности и ресурсосбережения коммерческого потенциала для работы «Разработка алгоритма автоматического восстановления позы скелета руки на основе карт глубины» были определены потенциальные потребители разработки – IT-компании, занимающиеся разработкой человеко-компьютерных интерфейсов, а также компании, работающие в области виртуальной и дополненной реальности.

Также был произведен анализ конкурентных технических решений, результаты которого показали, что предложенный в работе метод имеет высокую конкурентоспособность. Рассчитанный интегральный показатель конкурентоспособности подтвердил перспективность разработки и проведения исследований в данной области.

Сильные и слабые стороны решения, его возможности и угрозы, а также корреляция этих показателей были определенных в ходе SWOT-анализа. Полученные показатели позволили определить направление развития разработки для достижения наибольшей востребованности среди целевой аудитории. Результаты SWOT-анализа показали, что необходимо направить дальнейшую разработку в сторону повышения скорости и качества детектирования.

Определение перечня этапов и работ в рамках проведения научного исследования позволило структурировать и упорядочить запланированные этапы, а также распределить ответственных исполнителей-участников проекта. Созданный перечень этапов и работ в рамках проведения научного исследования лег в основу структуры календарного плана-графика, необходимого для детального планирования времени выполнения определенного этапа научного исследования.

Исходя из полученных данных и проведенного анализа эффективности, можно сделать вывод, что предложенный вариант исполнения является наиболее эффективным с позиции ресурсоэффективности, поскольку его интегральные показатели ресурсоэффективности разработки и эффективности выше, чем у других рассмотренных вариантов. Стоимость разработки является не самой низкой, однако потенциал разработки намного выше конкурентных решений. Именно поэтому был выбран текущий вариант исполнения.

5 СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ

Введение. Целью работы является разработка алгоритма автоматического восстановления позы скелета руки на основе карт глубины. В работе рассматриваются как алгоритмические способы решения поставленной задачи, так и нейросетевые методы искусственного интеллекта. К области практического применения данной разработки можно отнести реализацию пользовательских интерфейсов для устройств виртуальной и дополненной реальности; создание качественного человеко-компьютерного интерфейса, в частности, для людей с ограниченными возможностями.

Работы проводятся в учебно-научной лаборатории, далее офис, находящейся в аудитории 401 десятого корпуса Томского Политехнического Университета. Предоставленное рабочее место оборудовано столом, креслом, ЖК монитором и системным блоком компьютера.

Первичным этапом в задаче обеспечения безопасности труда является выявление возможных причин потенциальных несчастных случаев, производственных травм, профессиональных заболеваний, аварий и пожаров. Дальнейшими этапами являются разработка мероприятий по устранению выявленных причин и их реализация. Потенциальные причины и риски, а также конкретный набор мероприятий по их устранению, определяются спецификой выполняемых работ и априорными условиями труда (в частности, видом и состоянием рабочих мест исполнителей) [19].

5.1 Производственная безопасность

Производственные условия на рабочем месте характеризуются наличием различных опасных и вредных производственных факторов, оказывающих негативное влияние на работников. Под вредными факторами, понимают такие факторы трудового процесса и рабочей среды, которые характеризуются потенциальной опасностью для здоровья, в частности способствуют развитию каких-либо заболеваний, приводят к повышенной утомляемости и снижению работоспособности. При этом вредные факторы проявляются при

определенных условиях, таких как интенсивность и длительность воздействия. Опасные производственные факторы способны моментально оказать влияние на здоровье работника: привести к травмам, ожогам или к резкому ухудшению здоровья работников в результате отравления или облучения [18].

Выявленные опасные и вредные факторы согласно ГОСТ 12.0.003-74 (с измен. №1, октябрь 1978г., переиздание 1999 г.) «Классификация вредных и опасных производственных факторов» [18] приведены в таблице 24.

Таблица 24 – Выявленные опасные и вредные производственные факторы согласно ГОСТ 12.0.003-74

Наименование фактора	Действия на организм человека
Отклонения показателей микроклимата	Негативное влияние на нервную систему человека
Монотонность труда	Нервно-психическая перегрузка
Повышенные уровни электромагнитного излучения	Утомляемость, повышение или понижение кровяного давления, замедление пульса, выпадение волос, ломкость ногтей, сердечно-сосудистые заболевания
Повышенный уровень шума	Нарушение работы слухового аппарата, нервно-психические
Повышенный или пониженный уровень освещенности	Ухудшение зрения, травматизм
Повышенное значение напряжения в электрической цепи, замыкание которой может произойти через тело человека	Электрический удар, электротравма
Опасность возникновения пожара	Ожоги

Сегодня в стране более 50% лиц, работающих на ПЭВМ (персональной электронной вычислительной машине) или совмещающих эту работу с другой, имеют функциональные отклонения в состоянии здоровья. Защита от отрицательного воздействия ПЭВМ является одной из важнейших медико-биологических и социальных задач.

5.2 Анализ выявленных вредных факторов при разработке и эксплуатации проектируемого решения

Отклонение показателей микроклимата в помещении

Одним из необходимых благоприятных условий труда является обеспечение в помещениях нормальных метеорологических условий, оказывающих существенное влияние на самочувствие человека. Метеорологические условия в производственных помещениях (микроклимат), зависят от ряда особенностей технологического процесса, а также внешних условий (климата, сезона, условий вентиляции и отопления) [20].

Для безопасной работы необходимо соблюдать показатели микроклимата. Ниже приведены оптимальные величины показателей микроклимата на рабочих местах производственных помещений по СанПиН 2.2.4.548-96 (табл. 25). Работа программиста относится к категории 1а, потому что уровень энергозатрат не превышает 139 Вт.

Таблица 25 – Допустимые параметры микроклимата на рабочем месте

Период года	Категория работы	Температура воздуха, °С	Относительная влажность воздуха, %	Скорость движения воздуха, м/с
Холодный	1а	20 - 25	15 - 75	0,1
Теплый	1а	21 - 28	15 - 75	0,1

Оптимальные значения перечисленных параметров для работ с ПК, установленные санитарными нормами, приведены в таблице 26.

Таблица 26 – Оптимальные значения показателей микроклимата

Период года	Температура воздуха, °С	Относительная влажность воздуха, %	Скорость движения
Холодный	22-24	40-60	0,1
Теплый	23-25	40-60	0,1

Для создания благоприятных условий труда и повышения производительности, необходимо поддерживать оптимальные параметры микроклимата производственных помещений. Для этого предусмотрены

следующие средства: центральное отопление, вентиляция (искусственная и естественная), искусственное кондиционирование. Все из вышеперечисленных средств используются на рабочем месте.

Недостаточная освещенность рабочей зоны

Для производственных помещений всех назначений применяются системы общего равномерного освещения.

Источником света для общего освещения выберем люминесцентные лампы белой цветности (ЛБ). Мощность (ЛБ) 40 Вт. Световой поток 2800 лм.

Для помещения выберем тип светильника для люминесцентных ламп ШОД - 2-40. Количество ламп – 2. Мощность лампы 40 Вт. Длина светильника 1228 мм. Ширина светильника 284 мм.

Для создания благоприятных зрительных условий, наименьшая допустимая высота подвеса светильника ШОД над полом 2,5 м. План рабочего места инженера показан на рисунке 23.

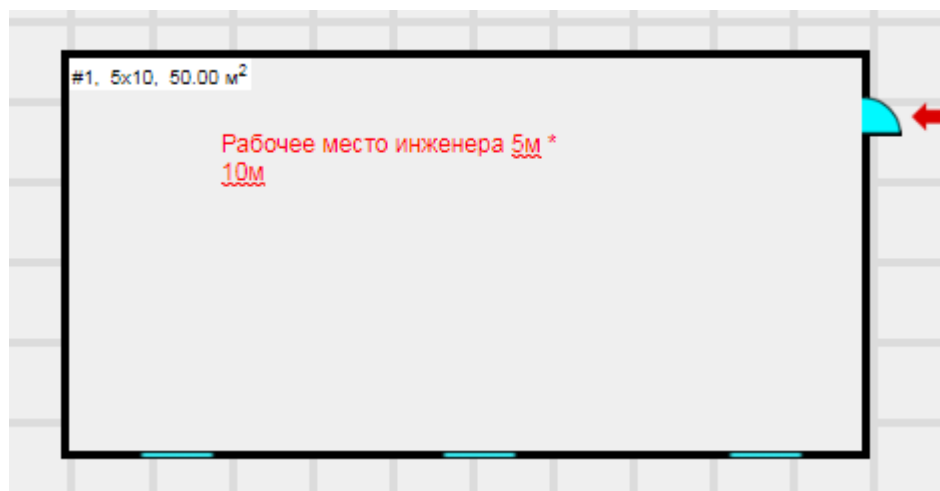


Рисунок 23 – Рабочее помещение инженера

Параметры рабочего места инженера: $A = 10\text{м}$ – длина помещения; $B = 5\text{м}$ – ширина помещения; $H = 3\text{м}$ – высота помещения; $S = 50\text{м}^2$ – площадь помещения.

Размещение светильников в помещении определяется следующими параметрами:

$h_c = 0,2\text{м}$ – расстояние светильников от перекрытия (свес);

$h_n = H - h_c = 3 - 0,2 = 2,8\text{м}$ – высота светильника над полом, высота подвеса;

$h_{rp} = 0,8\text{м}$ – высота рабочей поверхности над полом;

$h = h_n - h_{rp} = 2,8 - 0,8 = 2\text{м}$ – расчётная высота, высота светильника над рабочей поверхностью.

Уровень необходимого освещения определяется степенью точности зрительных работ. В таблице 27 приведены требования к освещению помещений промышленных зданий согласно СП 52.13330.20 СНиП 23-05-95 [21].

Таблица 27 – Нормативные значения освещённости согласно СНиП 23-05-95 [21]

Характеристика зрительной работы	Наименьший или эквивалентный размер объекта различения, мм	Разряд зрительной работы	Подразряд зрительной работы	Контраст объекта с фоном	Характеристика фона	Искусственное освещение		
						Освещённость, Лк		
						При комб. освещении		При общ.освещении
						всего	В том числе от общего	
Высокой точности	От 0,3 до 0,5	III	Г	Средний << Большой	Светлый << Средний	400	200	200

Индекс помещения находится по формуле:

$$i = \frac{S}{h \cdot (A + B)}$$

Индекс помещения $i = 1,667$

Коэффициент отражения потолка $R_n = 70\%$,

Коэффициент отражения стен $R_c = 50\%$

По индексу помещения и коэффициентам отражения определяем ближайший коэффициент использования светового потока $\eta = 0,48$.

Размещаем светильники в два ряда. В каждом ряду можно установить 6 светильников ШОД-2-40. Учитывая, что в каждом светильнике установлено две лампы белой цветности (ЛБ). Мощность (ЛБ) 40 Вт. Световой поток 2800 лм., общее число ламп в помещении $N = 24$.

По формуле 6.2 определяем световой поток лампы:

$$\Phi = \frac{E_n \cdot S \cdot K_z \cdot Z}{N \cdot \eta}$$

где Φ – световой поток лампы;

E_n – норма освещенности, 400 Лк;

S – площадь помещения, 50 м²;

K_z – коэффициент запаса, 1,5 – для люминесцентных ламп;

Z – коэффициент неравномерности освещения, 1,1 – для люминесцентных ламп;

N – число ламп в помещении, 24;

η – коэффициент использования светового потока, 0,48;

$$\Phi = \frac{400 \cdot 50 \cdot 1,5 \cdot 1,1}{24 \cdot 0,48} = 2865$$

Делаем проверку выполнения условия:

$$-10\% \leq \frac{\Phi_{\text{расч}} - \Phi_{\text{лампы}}}{\Phi_{\text{расч}}} \cdot 100\% \leq +20\%$$

$$\frac{2865 - 2800}{2865} \cdot 100\% = 2,27\%$$

Определяем электрическую мощность осветительной установки:

$$P = 24 \cdot 40 = 960 \text{ Вт}$$

Оптимальное расстояние L от крайнего ряда светильников до стены рекомендуется принимать равным $L/3$.

L – расстояние между соседними светильниками или рядами.

План размещения светильников в рабочем помещении показан на рисунке 24.

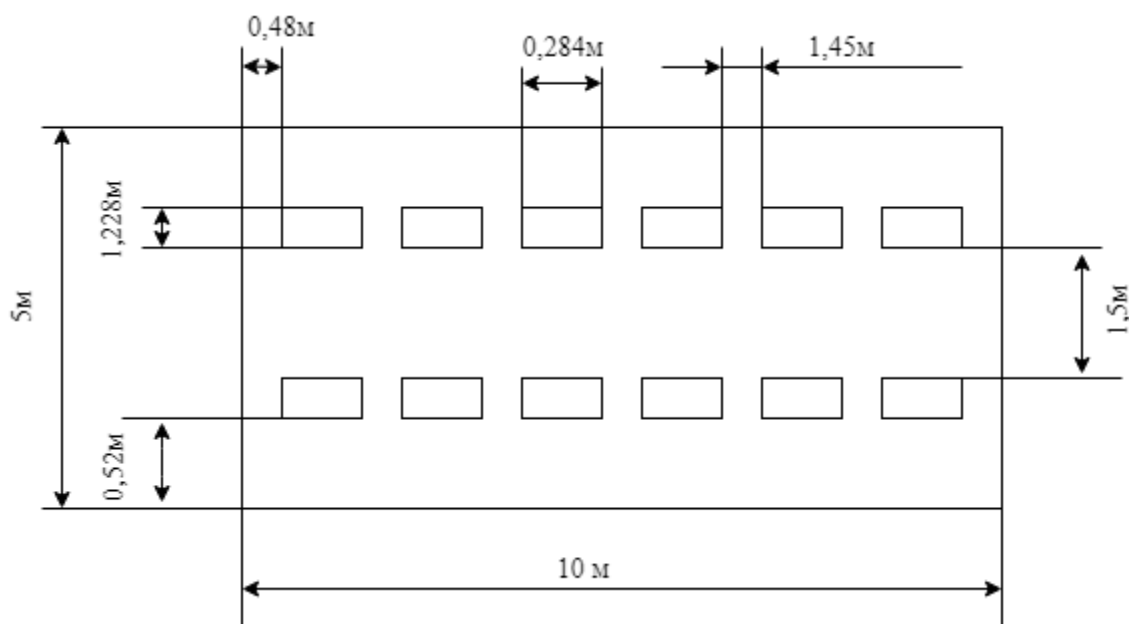


Рисунок 24 – План размещения светильников в помещении

Повышенный уровень шума и вибрации на рабочем месте

Шум определяется как звук, оцениваемый негативно и наносящий вред здоровью [19]. Основными источниками шума в помещении являются:

- система охлаждения центральных процессоров;
- жесткие диски.

При выполнении основной работы на ПЭВМ уровень шума на рабочем месте не должен превышать 50 дБА. Допустимые уровни звукового давления в помещениях для персонала, осуществляющего эксплуатацию ЭВМ при разных значениях частот согласно СН 2.2.4/2.1.8.562-96 «Шум на рабочих местах, в помещениях жилых, общественных зданий и на территории жилой, застройки», приведены в таблице 28 [23].

Таблица 28 – Допустимые уровни звука на рабочем месте согласно СН 2.2.4/2.1.8.562-96 [23]

Вид трудовой деятельности, рабочее место	Уровни звукового давления, дБ, в октавных полосах со среднегеометрическими частотами, Гц									Уровни звука и эквивалентного звука (в дБА)
	31,5	63	125	250	500	1000	2000	4000	8000	
Конструкторские бюро, программисты, лаборатории	86	71	61	54	49	45	42	40	38	50

Для снижения уровня шума, производимого персональными компьютерами, регулярно проводилось техническое обслуживание:

- чистка от пыли;
- замена смазывающих веществ;
- применение звукопоглощающих материалов.

Монотонность труда

Согласно ПОТ РО 29-002-94. «Правила по охране труда для издательств» [24] вредные психофизиологические подразделяются на:

- физические нагрузки (статические и динамические); умственное перенапряжение;
- перенапряжение анализаторов (глаз и др.); монотонность труда;
- эмоциональные перегрузки.

Для снижения влияние вредных и опасных психофизиологических факторов, влияющих на человека, основная деятельность которого предполагает работу за персональным компьютером, было осуществлено [24]:

- размещение персонального компьютера (ПК) на рабочем месте необходимо с учётом границ полей зрения работающего;
- использование рабочего стула с подъемно-поворотным устройством, обеспечивающим регулировку высоты сиденья и спинки, изменение угла наклона спинки;

- использование рабочего стула с подлокотниками. Высота поверхности сиденья регулировалась в пределах 400 – 500 мм. Ширина сиденья составляла – 400 мм, а глубина – 380 мм;
- использование рабочего стула с высотой опорной поверхности спинки – 300 мм, шириной 380 мм, углом наклона спинки в пределах 90 – 100° к плоскости сиденья;
- обеспечение оптимального режима труда и отдыха в течение рабочего дня;
- назначение частых, но коротких перерывов.

Повышенный уровень электромагнитных излучений

Воздействие электромагнитного излучения на человека зависит от напряженностей электрического и магнитного полей, потока энергии, частоты колебаний, размера облучаемого тела [19].

Работа проводилась на современном компьютере, где значения электромагнитного излучения малы и отвечают требованиям СанПиН 2.2.4.1191-03. «Электромагнитные поля в производственных условиях», которые приведены в таблице 29 [32].

Таблица 29 – Допустимые уровни электромагнитных полей согласно СанПиН 2.2.4.1191-03 [19]

Наименование параметров	Допустимые значения
Напряженность электромагнитного поля на расстоянии 50 см. вокруг ПЭВМ по электрической составляющей должна быть не более: • в диапазоне частот 5 Гц – 2 кГц • в диапазоне частот 2 – 400 кГц	25 В/м 2.5 В/м
Плотность магнитного потока должна быть не более: • в диапазоне частот 5 Гц – 2 кГц • в диапазоне частот 2 – 400 кГц	250 нТл 25 нТл
Напряженность электростатического поля:	20 кВ/м

Основной способ снижения вредного воздействия – это увеличение расстояния от источника (не менее 50 см от пользователя).

5.3 Анализ выявленных опасных факторов при разработке и эксплуатации проектируемого решения

5.3.1 Электробезопасность

Одним из выявленным опасных факторов является поражение электрическим током, так как напряжение считается безопасным при $U < 42\text{В}$, а вычислительная техника питается от сети 220 В с частотой 50 Гц. Ток является опасным, так как ток с частотой 20-100 Гц наиболее опасен. Результатом воздействия на организм человека электрического тока могут быть электрические травмы, электрические удары, и даже смерть.

Виды электротравм: местные электротравмы, к ним относятся: электрический ожог, электрические знаки, металлизация кожи, механические повреждения). Особую опасность представляют электрические травмы, которые выглядят в виде ожогов.

Электрический ожог возникает на том месте тела человека, в котором контакт происходит с токоведущей частью электроустановки. Электроожоги сопровождаются кровотечениями, омертвлением отдельных участков тела. Лечатся они гораздо труднее и медленнее обычных термических.

В результате механического повреждения могут разорваться кровеносные сосуды, нервные ткани, а также случаются вывихи суставов и даже переломы костей. Такие повреждения могут возникнуть в результате сокращений мышц под действием тока, который проходит через тело человека.

Электрические знаки в основном безболезненны, они могут возникнуть у 20% пострадавших от тока. Иногда электрические знаки выглядят в виде царапин, ушибов, бородавок, мозолей, также они представляют собой серые или бледно-желтые пятна круглоовальной формы с углублением в центре [25].

Чтобы защититься от поражения током, необходимо:

- обеспечить недоступность токоведущих частей от случайных прикосновений;
- электрическое разделение цепи;
- устранять опасности поражения при проявлении напряжения на разных частях.

При работе с компьютером при прикосновении к его элементам могут возникнуть токи статического электричества, которые в свою очередь имеют свойство притягивать пыль и мелкие частицы к экрану. Пыль на экране ухудшает видимость, а при подвижности воздуха может попасть на кожу лица и в легкие, что вызывает заболевание кожи и дыхательных путей.

Для предотвращения этого существуют специальные шнуры питания с заземлением и экраны для снятия статического электричества, это поможет защититься от статического электричества, а также необходимо проводить регулярную влажную уборку рабочего помещения.

По электробезопасности рабочее место относится к помещениям без повышенной опасности поражения людей электрическим током

характеризуются отсутствием условий, создающих повышенную или особую опасность. К ним относятся жилые помещения, лаборатории, конструкторские бюро, заводоуправление, конторские помещения и другие [27].

Степень воздействия зависит от продолжительности работы и индивидуальных особенностей организма.

Для снижения уровня воздействия, необходимо:

- экранирование экрана монитора;
- соблюдать оптимально расстояние от экрана;
- рационально размещать оборудование (если имеется несколько компьютеров, то расстояние между боковыми и задними стенками компьютеров должно быть 1,22 м);
- организовывать перерывы 10-15 минут через каждые 45-60 минут работы.

5.3.2 Пожаробезопасность

Одними из наиболее вероятных и разрушительных видов ЧС являются пожар или взрыв на рабочем месте. Пожарная безопасность представляет собой единый комплекс организационных, технических, режимных и эксплуатационных мероприятий по предупреждению пожаров и взрывов [26].

Причинами возгораний в рабочей зоне являются:

- резкие перепады напряжения;
- короткое замыкание в проводке, когда рубильник не отключен;
- пожар в соседнем помещении;
- короткое замыкание в розетке.

5.4 Экологическая безопасность

5.4.1 Анализ влияния процесса исследования на окружающую среду

Помещение с персональным компьютером относится к пятому классу, размер санитарно-защитной зоны которого равен 50 метров, так как работа на персональном компьютере не является экологически опасной.

5.4.2 Обоснование мероприятий по защите окружающей среды

Отработанная техника утилизируется с уничтожением информации согласно ГОСТ Р 50739-95 «Средства вычислительной техники. Защита от несанкционированного доступа к информации» [28]. Отработанные люминесцентные лампы утилизируются согласно ГОСТ 30772-2001 «Ресурсосбережение. Обращение с отходами. Термины и определения» [29].

В случае выхода из строя они списываются и отправляются на специальный склад, который при необходимости принимает меры по утилизации списанной техники и комплектующих.

В целом, утилизация ламп предполагает передачу использованных ламп предприятиям – переработчикам, которые с помощью специального оборудования перерабатывают вредные лампы в безвредное сырье – сорбент, которое в последующем используют в качестве материала для производства, например тротуарной плитки.

Утилизация макулатуры должна производиться в соответствии с ГОСТ Р 55090-2012 «Ресурсосбережение. Обращение с отходами. Рекомендации по утилизации отходов бумаги» [31] и ГОСТ 10700-97 «Макулатура бумажная и картонная. Технические условия» [32], согласно которым:

- макулатура должна разделяться на три группы: А – высокого качества, Б – среднего качества, В – низкого качества;
- макулатура каждой группы в зависимости от состава, источников поступления, цвета и способности к роспуску должна соответствовать маркам;
- макулатура должна содержать не более определённого количества примесей, определённого для каждой группы.

Все мероприятия по утилизации выполняются посредством сторонних компаний, имеющих соответствующие лицензии на такую деятельность (согласно ст. 17 Федерального закона от 8 августа 2001 г. №128-ФЗ «О лицензировании отдельных видов деятельности» [30]).

5.5 Защита в чрезвычайных ситуациях

5.5.1 Анализ вероятных ЧС, которые могут возникнуть на рабочем месте при проведении исследований

В данном случае на объекте (офис) могут возникать чрезвычайные ситуации (ЧС) следующего характера:

- техногенные;
- экологические;
- природные.

Наиболее типичной ЧС для помещения, котором производится выполнение ВКР, является пожар. Данная ЧС может произойти в случае замыкания электропроводки оборудования, обрыва проводов, не соблюдения мер пожаробезопасности и т.д.

5.5.2 Обоснование мероприятий по предотвращению ЧС и разработка порядка действия в случае возникновения ЧС

Пожарная безопасность включает в себя комплекс организационных и технических мероприятий, направленных на обеспечение безопасности людей, предотвращения пожара, ограничение его распространения, а также создание условия для успешного тушения пожара.

Пожарная опасность персональных электронно-вычислительных машин, обусловлена наличием в применяемом электрооборудовании горючих изоляционных материалов. Горючими являются изоляция обмоток трансформаторов, различных электромагнитов, проводов и кабелей.

Помещение, где проводились описанные выше работы, по пожарной и взрывной опасности относят к категории Д (пониженная пожароопасность) согласно ППБ-03 [25], характеризующейся отсутствием легковоспламеняющихся веществ и материалов в горячем состоянии. К

категории Д относятся помещения, в которых находятся (обращаются) негорючие вещества и материалы в холодном состоянии.

Здание 10-ого корпуса, в котором находится помещение, относится к негорючим по НПБ 105-95. «Определение категорий помещений и зданий по взрывопожарной и пожарной опасности» [26].

Для того что бы избежать возникновения пожара необходимо проводить следующие профилактические работы, направленные на устранение возможных источников возникновения пожара:

- периодическая проверка проводки;
- отключение оборудования при покидании рабочего места;
- проведение инструктажа работников о пожаробезопасности.
- Для предотвращения пожара в аудитории с ПЭВМ имеется:
- углекислотный огнетушитель типа ОУ-2 (данный тип огнетушителя подходит для помещений с электрооборудованием (ГОСТ Р 51057-01 и НПБ 155-02);
- пожарная сигнализация ДИП-3СУ (извещатель пожарный, дымовой оптико-электронный точечный).

В случае возникновения ЧС как пожар, необходимо предпринять меры по эвакуации персонала из офисного помещения в соответствии с планом эвакуации (план эвакуации имеется и размещен в помещении лаборатории). При отсутствии прямых угроз здоровью и жизни произвести попытку тушения возникшего возгорания огнетушителем. В случае потери контроля над пожаром, необходимо эвакуироваться вслед за сотрудниками по плану эвакуации и ждать приезда специалистов, пожарников. При возникновении пожара должна сработать система пожаротушения, издав предупредительные сигналы, и передав на пункт пожарной станции сигнал о ЧС, в случае если система не сработала, по каким либо причинам, необходимо самостоятельно произвести вызов пожарной службы по телефону 112, сообщить место возникновения ЧС и ожидать приезда специалистов.

5.6 Правовые и организационные вопросы обеспечения безопасности

5.6.1 Правовые нормы трудового законодательства

Продолжительность рабочего дня не должна превышать 40 часов в неделю. Возможно сокращение рабочего времени. Для работников, возраст которых меньше 16 лет – не более 24 часа в неделю, от 16 до 18 лет – не более 35 часов, как и для инвалидов I и II группы. Также рабочее время зависит от условий труда: для работников, работающих на рабочих местах с вредными условиями для жизни – не больше 36 часов в неделю

Светопроемы должны быть ориентированы преимущественно на север и северо-восток. В рабочем кабинете имеется три окна, два ориентированные на север и одно на северо-запад.

Площадь на одно рабочее место с ПЭВМ для взрослых пользователей должна составлять не менее 6,0 м².

Требования к организации и оборудованию рабочих мест с ПЭВМ по СанПиН 2.2.2/2.4.1340–03 представлены в таблице 30.

Таблица 30 – Требования к организации и оборудованию рабочих мест с ПЭВМ

Требование	Требуемые значения параметров	Значения параметров в комнате
Высота рабочей поверхности стола	680 – 800 мм	740 мм
Расположение монитора от глаз пользователя	600 – 700 мм	640 мм
Расположение клавиатуры на поверхности стола от края	100 – 300 мм	190 мм
Высота стула над полом (для роста 161–170 см)	420 мм	420 мм
Угол наклона монитора	0 – 30 градусов	5 градусов

5.6.2 Организационные мероприятия при компоновке рабочей зоны исследователя

Размер рабочей зоны, приходящейся на одного человека должен составлять не менее 4,5 м² (норма, указанная в СанПиН 2.2.2/2.4.1340-03).

Размер и форма рабочего стола должна позволять поместить все необходимое рабочее оборудование, а также располагать свободным местом для размещения документов и ведения рабочих записей.

Рекомендуемый рабочий стол, отвечающий основным параметрам (рис. 25).

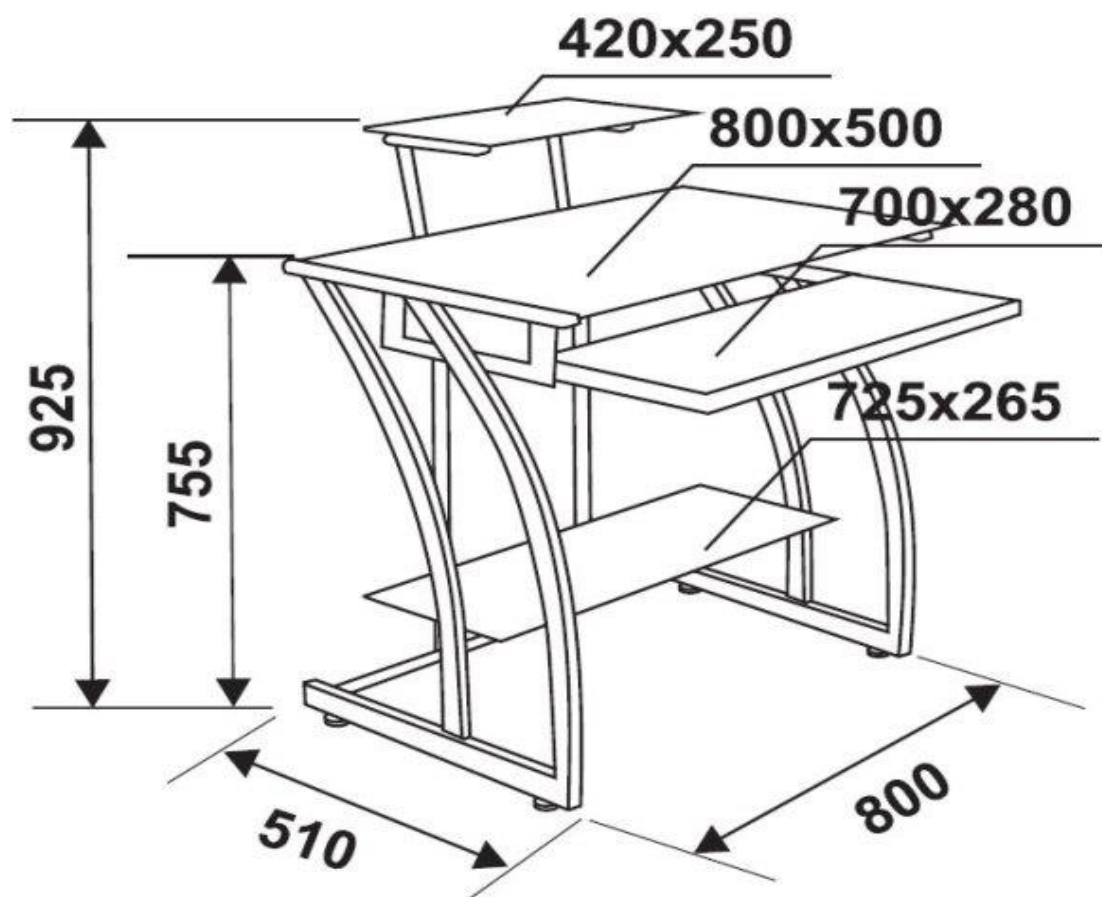


Рисунок 25 – Конструкция стола

Рабочее кресло должно состоять из следующих элементов: сиденье, спинка и подлокотники. Также могут быть предусмотрены подголовник и подставка для ног.

Рекомендуемый рабочий стул, отвечающий основным параметрам (рис. 26).

CH-213AXX

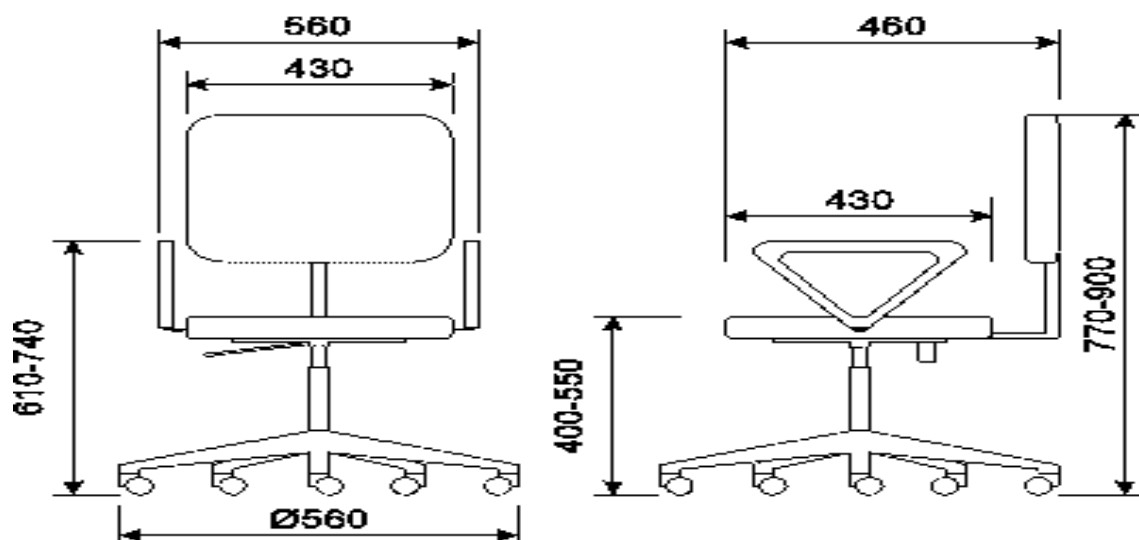


Рисунок 26 – Конструкция стула

Рабочее место также может быть оборудовано подставкой для ног со следующими параметрами: ширина не менее 300 мм, глубина не менее 400 мм, регулировка по высоте в пределах до 150 мм и угол наклона опорной поверхности подставки до 20°. Требования к поверхности подставки: рифленая и оснащенная бортиком высотой 10 мм.

Клавиатура располагается на расстоянии 100 - 300 мм от края, со стороны пользователя на специальной рабочей поверхности, регулируемой по высоте подставке, не находящейся на основной столешнице.

Расположение монитора и позиционирование пользователя (рис. 27).

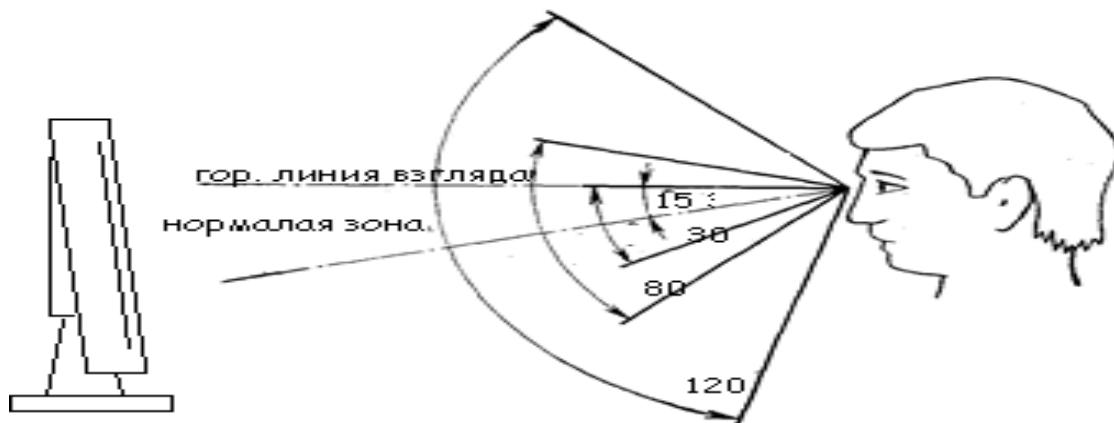


Рисунок 27 – Угол зрения

Заключение. В данной части работы изложены основные требования к рабочему месту инженера-программиста. Они должны создавать условия для комфортной работы. В соответствии с разработанными регламентами были предложены оптимальные размеры рабочего стола и кресла, рабочей поверхности, предложена система освещения рабочего помещения. Выполнение условий, необходимых для оптимальной организации рабочего места инженера-программиста, позволяет в течение всего рабочего дня сохранять высокую работоспособность, повышать производительность труда работника.

В части экологии производителям и потребителям компьютерной техники необходимо серьезно задуматься о снижении вредного воздействия на окружающую среду посредством разработки регламентирующих документов по утилизации данного оборудования, отправки его в переработку, для возможного дальнейшего вторичного использования.

ЗАКЛЮЧЕНИЕ

В ходе дипломной работы был выполнен аналитический обзор общедоступных методов обработки карт глубины. Был выполнен обзор научной литературы, проанализированы представленные в литературе подходы к задаче определения позиции скелета руки с использованием 2D и 3D данных.

На основании проанализированной информации были предложены два подхода к решению поставленной задачи: алгоритмический подход с использованием алгоритма ICP и подход с использованием сверточных нейронных сетей.

В соответствии с предложенной архитектурой были разработаны и реализованы предложенные методы, позволяющие на основе карт глубины восстановить местоположения суставов руки.

Реализованные методы были протестированы как на общедоступном тестовом наборе, так и на собственном наборе изображений. Одним из главных достоинств разработанного нейросетевого метода является использование нескольких проекций облака точек, что позволяет повысить точность детектирования положения суставов.

Полученные результаты свидетельствуют о том, что разработанный нейросетевой метод имеет преимущество над методами, работающими с одной проекцией или непосредственно с картой глубины.

Запланированным развитием работы является оптимизация предложенного нейросетевого метода с целью увеличения производительности, а также повышение точности детектирования.

СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

1. Скелетная анимация, прямая и инверсная кинематика [Электронный ресурс] – Режим доступа: <http://www.mir3d.ru/learning/974> (дата обращения: 10.12.2018).
2. Официальный сайт Node.js. // Платформа Node.js. [Электронный ресурс] – Режим доступа: <https://nodejs.org/en/docs/> (дата обращения: 17.12.2018).
3. James S. Supancic, Grégory Rogez, Yi Yang, Jamie Shotton, Deva Ramanan. Depth-Based Hand Pose Estimation: Data, Methods, and Challenges. ICCV 2015.
4. Iason Oikonomidis, Nikolaos Kyriazis, Antonis A. Argyros. Efficient Model-based 3D Tracking of Hand Articulations using Kinect. BMVC 2011.
5. Hui Liang, Shunfeng Wang, Qian Sun, Yong-Jin Liu, Junsong Yuan, Jun Luo, Ying He. Barehanded music: real-time hand interaction for virtual piano. I3D 2016.
6. Chi Xu, Li Cheng. Efficient Hand Pose Estimation from a Single Depth Image. ICCV 2013.
7. Danhang Tang, Tsz-Ho Yu, Tae-Kyun Kim. Real-Time Articulated Hand Pose Estimation Using Semi-supervised Transductive Regression Forests. ICCV 2013.
8. Furkan Kırac, Yunus Emre Kara, Lale Akarun. Hierarchically constrained 3D hand pose estimation using regression forests from single frame depth data. Pattern Recognition Letters 2014.
9. Alexander Toshev, Christian Szegedy. DeepPose: Human Pose Estimation via Deep Neural Networks. CVPR 2014.
10. Xiao Sun, Yichen Wei, Shuang Liang, Xiaoou Tang, Jian Sun. Cascaded hand pose regression. CVPR 2015.

11. Arjun Jain, Jonathan Tompson, Mykhaylo Andriluka, Graham W. Taylor, Christoph Bregler. Learning Human Pose Estimation Features with Convolutional Networks. ICLR 2014.
12. Jonathan Tompson, Ross Goroshin, Arjun Jain, Yann LeCun, Christopher Bregler. Efficient Object Localization Using Convolutional Networks. CVPR 2015.
13. Jonathan Tompson, Murphy Martin Stein, Yann Lecun, Ken Perlin. Real-Time Continuous Pose Recovery of Human Hands Using Convolutional Networks. ACM Transactions on Graphics 2014.
14. Chen Qian, Xiao Sun, Yichen Wei, Xiaoou Tang, Jian Sun. Realtime and Robust Hand Tracking from Depth. CVPR 2014.
15. OpenPose – система для определения ключевых точек человеческого тела, рук, ног и лиц [Электронный ресурс] – Режим доступа: <https://github.com/CMU-Perceptual-Computing-Lab/openpose> (дата обращения: 04.03.2019).
16. Keras – открытая нейросетевая библиотека, написанная на языке Python [Электронный ресурс] – Режим доступа: <https://keras.io> (дата обращения: 11.02.2019).
17. СанПиН 2.2.2/2.4.1340-03. «Гигиенические требования к персональным электронно-вычислительным машинам и организации работы».
18. ГОСТ 12.0.003-74 (с измен. №1, октябрь 1978г., переиздание 1999 г.) «Классификация вредных и опасных производственных факторов».
19. Безопасность жизнедеятельности: Учебник для вузов /Под ред. К.З. Ушакова. – М.: Изд-во Московского гос. горного университета, 2000.– 430 с.
20. СанПиН 2.2.4.548 – 96. «Гигиенические требования к микроклимату производственных помещений».
21. СНиП 23-05-95. «Естественное и искусственное освещение».
22. СНиП 23-05-2010 «Нормы Коэффициента пульсации помещений административных зданий».

23. СН 2.2.4/2.1.8.562-96 «Шум на рабочих местах, в помещениях жилых, общественных зданий и на территории жилой застройки».
24. ПОТ РО 29-002-94. «Правила по охране труда для издательств».
25. ГОСТ 12.1.009-76 «Электробезопасность. Термины и определения».
26. СНиП 21-01-97. «Пожарная безопасность зданий и сооружений».
27. НПБ 105-95. «Определение категорий помещений и зданий по взрывопожарной и пожарной опасности».
28. ГОСТ Р 50739-95 «Средства вычислительной техники. Защита от несанкционированного доступа к информации».
29. ГОСТ 30772-2001 «Ресурсосбережение. Обращение с отходами. Термины и определения».
30. Федеральный закон от 8 августа 2001 г. №128-ФЗ «О лицензировании отдельных видов деятельности».
31. ГОСТ Р 55090-2012 «Ресурсосбережение. Обращение с отходами. Рекомендации по утилизации отходов бумаги».
32. ГОСТ 10700-97 «Макулатура бумажная и картонная. Технические условия».
33. СанПиН 2.2.4.1191-03. «Электромагнитные поля в производственных условиях».
34. ГОСТ Р 56276-2014 «Газы парниковые. Углеродный след продукции. Требования и руководящие указания по количественному определению и предоставлению информации».
35. ГОСТ Р ИСО 14046-2017 «Экологический менеджмент. Водный след. Принципы, требования и руководящие указания».

ПРИЛОЖЕНИЕ А

(справочное)

OVERVIEW

Студент

Группа	ФИО	Подпись	Дата
8BM73	Иовлев Ю. Д.		

Консультант Отделения информационных технологий ИШИТР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ	Мирошниченко Е.А.	К.Т.Н.		

Консультант-лингвист отделения иностранных языков ШБИП

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИЯ	Сидоренко Т.В.	К.П.Н.		

1.1 Subject field description

Over the last few years, gesture estimation based on visual information has received much attention.

Increased interest in this area can be attributed to two majority factors: deep visualization and deep learning. First of all, hand gesture estimation methods have completely changed with the advent of depth sensors, as products like MS Kinect and Intel RealSense have become widespread and accessible.

As a source of 2.5D information, depth significantly resolves most of the ambiguities present in an RGB monocular. Secondly, deep learning has radically changed the approaches to solving computer vision problems.

The problem of 3D hand pose estimation has aroused a lot of attention in computer vision community for long, as it plays a significant role in human-computer interaction such as virtual or augmented reality applications. Despite the recent progress in this field, robust and accurate hand pose estimation remains a challenging task. Due to large pose variations and high dimension of hand motion, it is generally difficult to build an efficient mapping from image features to articulated hand pose parameters.

Data-driven methods for hand pose estimation train discriminative models, such as isometric self-organizing map, random forests and convolutional neural networks (CNNs), to map image features to hand pose parameters. With the availability of large annotated hand pose datasets, data-driven approaches become more advantageous as they do not require complex model calibration and are robust to poor initialization.

1.2 Literature review

Vision-based hand pose estimation has been extensively studied in literature over many years. The most common hand pose estimation techniques can be classified into model-driven approaches and data-driven approaches [3].

Model-driven methods usually find the optimal hand pose parameters via fitting a deformable 3D hand model to input image observations. Such methods have demonstrated to be quite effective, especially with the depth cameras [15]. However, there are some shortcomings for the model-driven methods. These shortcomings will be considered further.

The data-driven methods do not need the explicit specification of the hand size and motion constraints. Rather, such information is automatically encoded in the training data. Therefore, many recent methods are built upon such a scheme [5]. Among them, the random forest and its variants have proved to be reasonably accurate and fast. In [6], the authors propose to use the random forest to directly regress for the hand joint angles from depth images, in which a set of spatial-voting pixels cast their votes for hand pose independently and their votes are clustered into a set of candidates. The optimal one is determined by a verification stage with a hand model. A similar method is presented in [7], which further adopts transfer learning to make up for the inconsistency between synthesis and real-world data. As the estimations from random forest can be ambiguous for complex hand postures, pre-learned hand pose priors are sometimes utilized to better fuse independently predicted hand joint distributions [8].

Very recently, convolutional neural networks have shown to be effective in articulated pose estimation. In [9] they are tuned to regress for the 2D human poses by directly minimizing the pose estimation error on the training data. The results have shown to outperform the traditional methods largely. However, it takes more than twenty days to train the network and the dataset only contains several thousand images. Considering the relatively small size of the dataset used in [9], it can be

difficult to use it on larger datasets such as [10], which consist of more than 70000 images.

Also, it is reported in [11] that such direct mapping with CNNs from image features to continuous 2D/3D locations is of high nonlinearity and complexity as well as low generalization ability, which renders it difficult to train CNNs in such a manner.

To this end, in their work on body pose estimation [12], the CNNs are used to predict the heat-maps of joint positions instead of the original articulated pose parameters, and on each heat-map the intensity of a pixel indicates the likelihood for a joint occurring there. During training, the regression error is instead defined as the L2-norm of the difference between the estimated heat-map and the ground truth heat-map. y, the network can be trained efficiently and they achieve state-of-the-art performances.

Similarly, such a framework has also been applied in 3D hand pose estimation [13]. However, the heat-map only provides 2D information of the hand joint and the depth information is not fully utilized. To address this issue, a model-based verification stage is adopted to estimate the 3D hand pose based on the estimated heat-maps and the input depth image. Such heat-map based approaches are interesting as heat-maps can reflect the probability distribution of 3D hand joints in the projection plane

Inspired by such methods, heat-maps of multiple views are generated and fuse them together to estimate the probability distribution of hand joints in 3D space.

1.3 Theoretical basis

1.3.1 Depth maps

The depth map – a kind of image which is composed of gray pixels defined by values from 0 to 255 (fig. 1).

The "0" value of gray pixels stand for that "3D" pixels are located at the most distant place in the 3D scene while the "255" value of gray pixels stand for that "3D" pixels are located at the most near place. In depth map, each depth pixel would define the position in Z-axis where its corresponding 2D pixel will be located.

Depth maps are gaining popularity with the advent of cameras with a depth sensor. Depth maps are used in the following tasks:

- To create and generate autostereograms and in other related applications intended to create the illusion of 3D viewing through stereoscopy;
- In computer vision to model 3D shapes or reconstruct them;
- In machine vision and computer vision, to allow 3D images to be processed by 2D image tools.



Figure 128 – Depth map

1.3.2 Camera matrix

In computer vision a camera matrix or projection matrix is a 3×4 matrix, which describes the mapping of a pinhole camera from 3D points in the world to 2D points in an image.

The camera matrix is obtained by performing a camera calibration, there are tools that allow to perform camera calibration, both in manual mode and in automatic mode.

The 3D and 2D coordinates of the image are associated with the following formula $x = P \cdot X$, where x – 2D coordinates of the image, P – the camera matrix, X – the world coordinates of the point. Figure 2 shows the geometric representation of this transformation.

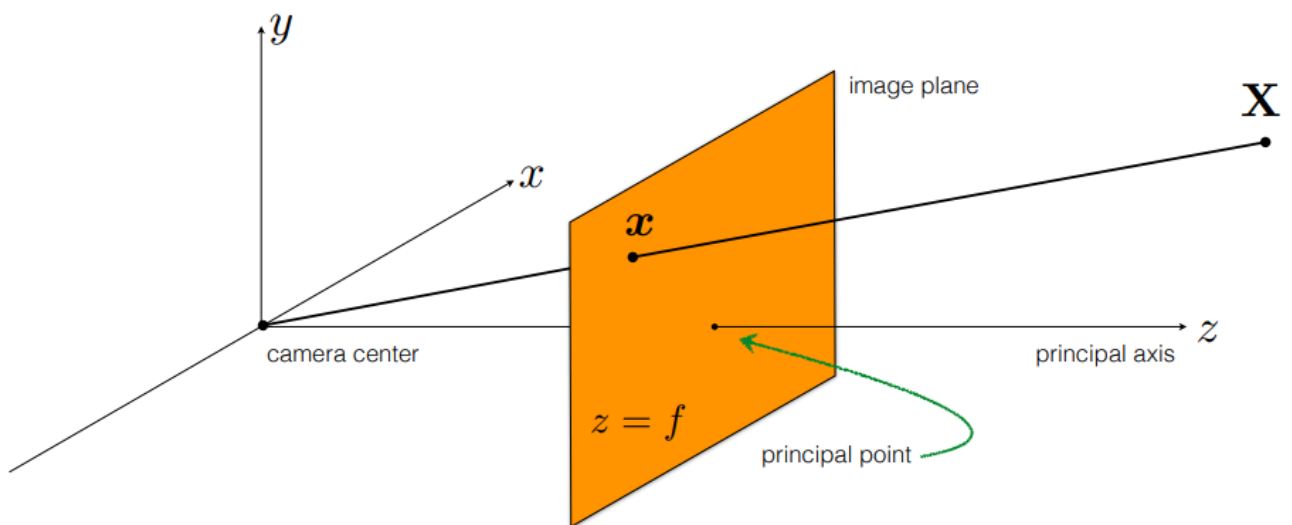


Figure 29 – Pinhole camera geometry

The figure shows that to perform coordinate transformations, need to define the focal distance of the camera and the principal point – the intersection of the camera principal axis with the image plane. These parameters are determined at the calibration stage.

Move to homogeneous coordinates, obtain the following relation:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} f & 0 & p_x & 0 \\ 0 & f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1)$$

The resulting formula is valid only if the origin of the world coordinate system coincides with the position of the camera. In the case when the camera is not at the beginning of the world coordinate system, it is required to determine the translation and rotation relative to the world coordinate system. This matrix is called the camera extrinsic and has the following form.

$$\begin{bmatrix} r1 & r2 & r3 & : & t1 \\ r4 & r5 & r6 & : & t2 \\ r7 & r8 & r9 & : & t3 \end{bmatrix}$$

The left side of the matrix is responsible for the rotation, and the right for the translation.

In this way, the camera matrix has the following form.

$$P = \begin{bmatrix} f & 0 & p_x & 0 \\ 0 & f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r1 & r2 & r3 & : & t1 \\ r4 & r5 & r6 & : & t2 \\ r7 & r8 & r9 & : & t3 \end{bmatrix} \quad (2)$$

With this matrix, can perform transformations from 2D coordinates to 3D and back.

1.3.3 Skeletal animation

Skeletal animation – a technique in computer animation.

Often the skeleton is a hierarchical set of interconnected bones, where each subsequent bone is connected with the previous one, that inherits all affine transformations by the hierarchy in the skeleton [1].

Further, all the vertices of the 3D model are connected to skeleton joints with the influence coefficient. In this way, when a joint is rotated, all the vertices of the 3D model attached to it also move. Due to this approach, the definition of skeletal animation is greatly simplified, because there is no need to animate each vertex of the model separately, and it is enough just to perform affine transformations over the skeletal joints.

An articulated joint is one of the main types of joints between the bones of the skeleton, such a connection is often used in the construction of skeletal models.

In articulations, the end point of one bone rotates around a static point, usually the endpoint of the parent bone. A feature of this relationship is that to move each bone to a new position, only rotation is specified. It follows that for animation it is necessary to store only information about the local rotation.

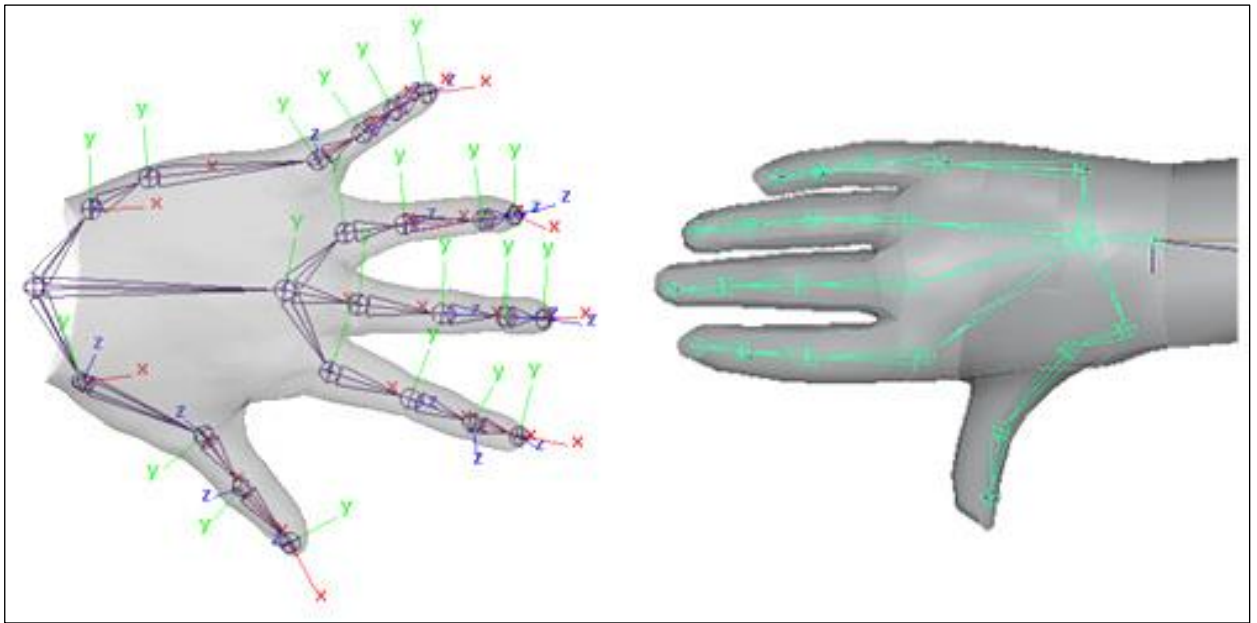


Figure 30 – Skeleton hand model

In this way, it is possible to define a simple way of representing the skeleton, in figure 3 on the left side, the skeleton of the arm is shown with the direction of the axes for each joint, and on the right side is the skeleton associated with the three-dimensional model. The set of bones has a hierarchical structure, each bone is defined by a local set of parameters – length and rotation. Rotation is usually specified by a quaternion or vector of rotation angles.

By convention, each bone is aligned along the X-axis. The endpoint of the bone is on the X-axis and is offset along with it by the length of the bone. The bone rotation matrix is calculated from the local rotation value of the current bone, and the bone translation is an offset by the length of the parent bone along the X-coordinate in the local coordinate system of the parent.

ПРИЛОЖЕНИЕ Б

CD-ДИСК