

На правах рукописи

**Козлова Юлия Ханифовна**

**МЕТОД СОЗДАНИЯ ПАРАМЕТРИЗОВАННОГО АВАТАРА ГОЛОВЫ ЧЕЛОВЕКА  
НА ОСНОВЕ НЕЙРОСЕТЕВОЙ МОДЕЛИ РЕНДЕРИНГА**

**2.3.8. – Информатика и информационные процессы**

**А в т о р е ф е р а т**  
диссертации на соискание ученой степени  
кандидата технических наук

Самара – 2024

Работа выполнена в федеральном государственном автономном образовательном учреждении высшего образования «Самарский национальный исследовательский университет имени академика С. П. Королева» на кафедре геоинформатики и информационной безопасности.

**Научный руководитель:**

**Мясников Владислав Валерьевич**, доктор физико-математических наук, профессор.

**Официальные оппоненты:**

**Дементьев Виталий Евгеньевич**, доктор технических наук, федеральное государственное бюджетное образовательное учреждение высшего образования «Ульяновский государственный технический университет», кафедра «Радиотехника, телекоммуникации и защита информации», заведующий кафедрой;

**Николаев Дмитрий Петрович**, доктор технических наук, федеральное государственное учреждение «Федеральный исследовательский центр «Информатика и управление» Российской академии наук», руководитель отдела зрительных систем 9.5.

**Ведущая организация:**

Федеральное автономное учреждение «Государственный научно-исследовательский институт авиационных систем», г. Москва.

Защита состоится 19 декабря 2024 года в 10:00 часов на заседании диссертационного совета 24.2.379.09, созданного на базе федерального государственного автономного образовательного учреждения высшего образования «Самарский национальный исследовательский университет имени академика С. П. Королева» по адресу: 443086, г. Самара, Московское шоссе, 34.

С диссертацией можно ознакомиться в библиотеке федерального государственного автономного образовательного учреждения высшего образования «Самарский национальный исследовательский университет имени академика С. П. Королева» и на официальном сайте по адресу: [https://ssau.ru/storage/pages/6527/file\\_66fc63119cbf69.83881780.pdf](https://ssau.ru/storage/pages/6527/file_66fc63119cbf69.83881780.pdf)

Автореферат разослан «\_\_\_» \_\_\_\_\_ 2024 г.

Учёный секретарь  
диссертационного совета

С. В. Востокин

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

### Актуальность темы исследования

Задача создания цифровых аватаров людей становится все более актуальной в последние годы в связи со стремительным развитием технологий виртуальной, смешанной и дополненной реальности. В областях киноиндустрии и игровой индустрии также существует высокий спрос на такие технологии. При этом для создания реалистичных аватаров требуется дорогостоящее оборудование, которое включает в себя специализированное освещение и многокамерную установку, а также кропотливый ручной труд специалистов, которые занимаются постобработкой полученных данных и проектированием примитивов для анимации, интегрируемых в существующие рендереры. Методы создания цифровых аватаров позволяют решать задачу телеприсутствия, что было особенно актуально во время пандемии COVID-19, которая в результате привела к частичному переходу на дистанционный формат работы, обучения, участия в различных мероприятиях, например, в научных конференциях. Wang T.-C., Mallya A. и Liu M.-Y. в своей работе показали, что интеграция технологии цифрового аватара в видеоконференции позволяет при снижении пропускной способности канала сохранять исходное качество изображения, так как не требуется выполнять передачу и сжатие всего изображения. Вместо этого на приемную сторону передается низкоразмерная закодированная информация о ключевых точках лица и положении головы, которая используется для синтеза реалистичной анимации. Ещё одним примером использования цифровых аватаров являются звонки с эффектом присутствия с помощью гарнитуры смешанной реальности. Так, в июне 2023 года компания Apple представила гарнитуру смешанной реальности «Apple Vision Pro», которая с помощью специализированного приложения «Persona» позволяет создать персонализированного аватара. Процесс создания аватара включает в себя процедуру сканирования головы камерами устройства в соответствии с заданными инструкциями. Конечное назначение представленной технологии – звонки с имитацией присутствия в едином физическом пространстве, где за счет информации с датчиков и камер гарнитур производится анимация аватаров.

Здесь и далее под **аватаром головы человека** понимается цифровое представление поверхности головы (форматами представления могут выступать полигональная сетка, облако точек, нейронное неявное представление, изображение и т. п.), полученное на основе некоторых данных (трехмерное представление головы, полученное в результате сканирования; изображение; набор изображений; видеопоследовательность и т. п.), которое может быть использовано для передачи и воспроизведения изображения лица/головы без потери идентичности. Под параметризованным аватаром будет пониматься аватар, для которого возможно выполнить синтез изображений, при котором мимика и поза головы будут управляться значениями некоторых параметров.

Таким образом, задача создания параметризованного аватара головы человека заключается в разработке метода, который принимает на вход некоторую информацию, описывающую внешний вид головы человека, а на выходе формирует реалистичное представление поверхности головы, которое может быть модифицировано в зависимости от ожидаемого выражения лица и/или положения головы, а также использовано для воспроизведения/реконструкции изображения/видеопоследовательности головы человека.

Для объективного анализа актуальности выбранной темы исследования была проведена агрегация статей с упоминанием ключевых слов «Head avatar», «NeRF», «Neural rendering и NeRF» по годам, начиная с 2015 года. В рассмотрении участвовали работы, размещенные в электронном архиве с открытым доступом для научных статей и рукописей arXiv.org. На рисунке 1 представлен результат анализа. Исходя из представленных результатов можно заметить нелинейный рост количества статей, посвященных как задаче создания аватаров головы человека, так и выбранному в рамках диссертационного исследования способу пространственного представления – Neural Radiance Fields (подробнее в подразделе 3.1.1 диссертации), в частности его модификации, где для рендеринга итогового изображения применяется не алгоритм объеметрического рендеринга (англ. volume rendering), а двумерная

свёрточная нейронная сеть, значительно ускоряющая эту процедуру. Использование NeRF в качестве способа представления поверхности предоставляет возможность для формирования трехмерной сцены с использованием разреженных данных.

Видно, что направление исследований, посвященное методам создания аватара головы человека, начало набирать популярность около пяти лет назад, однако значимых работ близких к теме диссертационного исследования от авторов из Российской Федерации или стран СНГ обнаружить не удалось.

Большое количество статей с высокими показателями цитируемости опубликовано авторами из Германии (Университет Макса Планка), США (Университет Стэнфорд, Университет Беркли, Университет Карнеги), Швейцарии (Швейцарская высшая техническая школа Цюриха), Великобритании (Оксфордский университет), Китая (Шанхайский университет) и других государств.

При анализе существующих решений по теме диссертационного исследования (см. первый раздел диссертации) были выявлены следующие недостатки:

1. Низкая доступность в связи с использованием сложных многокамерных установок (в некоторых случаях также требуется специализированное освещение).
2. Низкая степень схожести между субъектом и синтезируемыми изображениями.
3. Высокая вычислительная сложность на этапах обучения и/или синтеза новых изображений.

Перечисленные недостатки присущи рассмотренным методам в разной степени, например, некоторые методы обладают высокой доступностью и низкой вычислительной сложностью, но не способны синтезировать изображения с высокой степенью схожести, другие методы, напротив, способны синтезировать высококачественные фотореалистичные изображения, но требуют большого количества вычислений и/или сложных многокамерных систем.

Таким образом, ключевым направлением диссертационного исследования является разработка метода, который позволит устранить указанные недостатки. А именно, разрабатываемый метод создания аватара должен удовлетворять следующим условиям:

1. Доступность, под которой подразумевается возможность использования одной видеопоследовательности.
2. Высокая степень схожести между субъектом и синтезируемыми изображениями.
3. Низкая вычислительная сложность по сравнению с существующими решениями.

Учитывая изложенные тезисы, как тема диссертационной работы в целом, так и отдельные выбранные направления исследований в частности являются безусловно актуальными.

#### Цели и задачи исследования

Целью диссертационного исследования является разработка методов и алгоритмов создания параметризованного аватара головы человека, обеспечивающих при относительно низких вычислительных затратах процесса создания повышенное качество синтезируемых изображений-проекций.

Для достижения поставленной цели в диссертации решаются следующие задачи:

1. Разработка алгоритма оценки параметров модели головы FLAME с использованием RGBD изображения.
2. Разработка и исследование параметрической модели головы человека на основе нейросетевой модели представления поверхности CNeRF, двумерного нейронного рендеринга и синтетического набора данных, генерируемого в реальном времени.

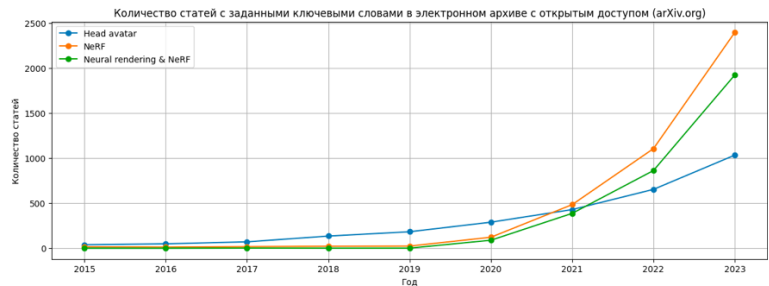


Рисунок 1 – Количество статей с заданными ключевыми словами по годам.

3. Разработка и исследование метода создания параметризованного аватара головы человека на основе разработанной параметрической модели головы человека.

Поставленные задачи определяют структуру работы и содержание ее разделов.

#### Методология и методы исследования

В диссертационной работе используются методы статистического анализа, компьютерной графики, компьютерного зрения и методы машинного обучения.

#### Научная новизна работы

1. Предложен алгоритм оценки параметров модели FLAME с использованием RGBD изображения.
2. Разработана параметрическая модель головы человека на основе нейросетевой модели представления поверхности CNeRF, двумерного нейронного рендеринга и синтетического набора данных, генерируемого в реальном времени.
3. Предложен метод создания параметризованного аватара головы человека на основе разработанной параметрической модели головы человека.
4. Предложен способ аугментации (расширения) реального набора данных, содержащего кадры видеопоследовательности с изображением головы человека, с использованием интерполяции промежуточных кадров.

#### Практическая значимость работы

Разработанные методы и алгоритмы могут быть использованы в составе систем телеприсутствия; виртуальной, дополненной, смешанной реальности; систем видеоконференцсвязи; систем создания видеоконтента и прочих системах, требующих синтеза визуально реалистичных трехмерных моделей лиц пользователей.

#### На защиту выносятся:

1. Алгоритм оценки параметров модели FLAME с использованием RGBD изображения, позволяющий достичь высокой точности трехмерной реконструкции.
2. Параметрическая модель головы человека на основе нейросетевой модели представления поверхности CNeRF, архитектуры сети двумерного нейронного рендеринга с блоком повышения пространственной дискретизации, позволяющим ускорить сходимость метода создания аватара, и синтетического набора данных, генерируемого в реальном времени.
3. Метод создания параметризованного аватара головы человека на основе разработанной параметрической модели головы человека, позволяющий достичь высокой скорости создания аватара конкретного человека, а также обеспечить высокую скорость синтеза изображений-проекций аватара при высоком значении показателя качества.
4. Способ аугментации (расширения) реального набора данных, содержащего кадры видеопоследовательности с изображением головы человека, с использованием интерполяции промежуточных кадров, который позволяет повысить качество синтезируемых изображений-проекций аватара для коротких видеопоследовательностей.

#### Соответствие специальности

Диссертация соответствует паспорту научной специальности 2.3.8. –Информатика и информационные процессы и охватывает следующие области исследования, входящие в эту специальность:

1. Разработка компьютерных методов и моделей описания, оценки и оптимизации информационных процессов и ресурсов, а также средств анализа и выявления закономерностей на основе обмена информацией пользователями и возможностей используемого программно-аппаратного обеспечения.
13. Разработка и применение методов распознавания образов, кластерного анализа, нейросетевых и нечетких технологий, решающих правил, мягких вычислений при анализе разнородной информации в базах данных.

### Степень достоверности и апробация результатов

Основные результаты диссертации были представлены на 3 научных конференциях:

1. Международной конференции «Информационные технологии и нанотехнологии» (ИТНТ, Самара, Россия) - 2022 год.
2. Международной конференции «Информационные технологии и нанотехнологии» (ИТНТ, Самара, Россия) - 2023 год.
3. Международной конференции «Информационные технологии и нанотехнологии» (ИТНТ, Самара, Россия) - 2024 год.

По теме диссертации опубликовано восемь работ [1–8]. Из них одна работа опубликована в изданиях, рекомендуемых ВАК, три работы опубликованы в рецензируемых изданиях, входящих в международные реферативные базы данных Scopus и/или WebOfScience. Пять работ выполнено без соавторов.

Результаты диссертационной работы:

1. Использованы в АО «Самара-Информспутник» при выполнении хозяйственных работ № 4/2021 от 22.09.2021 (2021–2023), № 7/2021 от 08.11.2021 (2021–2023).
2. Использованы в ФГУП «ГосНИИПП» в рамках составных частей научно-исследовательских работ по договорам №4/2021 от 22.09.2021 года и №7/2021 от 08.11.2021 года.
3. Использованы в рамках договора №69/12/2023 от 25.12.2023, в рамках гранта от Фонда Содействия Инновациям № 358ГС1ЦТС10-D5/91117 от 18.12.2023 в ООО «Давтех».
4. Использованы в учебном процессе в ФГАОУ ВО «Самарский национальный исследовательский университет имени академика С. П. Королева» в курсе лекций по дисциплине «Безопасность мультимедиа».

### Структура диссертации

Диссертация состоит из четырех разделов, заключения, списка литературы из 102 наименований; изложена на 210 страницах машинописного текста, содержит 59 рисунков, 2 таблицы, 2 приложения.

## **КРАТКОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИИ**

**В первом разделе** диссертационного исследования дана оценка современного состояния задачи создания цифрового аватара головы человека. Приведен обзор большого количества работ, представленных на различных научных конференциях высокого класса и опубликованных в журналах с высоким рангом. Производится систематизация этих работ и разработка системы классификации методов создания аватара головы человека.

Так, методы создания цифровых аватаров головы можно классифицировать по следующим критериям:

1. По способу представления поверхности цифрового аватара (явное представление поверхности; неявное представление поверхности; без информации о поверхности).
2. По обобщенности итоговой модели (общая модель; персональная модель).
3. По требованиям к набору данных для обучения (3D набор данных; 2D набор данных; синтетический набор данных).
4. По возможности управления параметрами модели («распутанное» пространство параметров; пространство без «распутывания» параметров).

Важно также отметить еще один аспект, который не является критерием классификации, но является важным при эксплуатации разработанного метода – необходимость предварительной оценки параметров параметрической модели. Данный этап подразумевает получение некоторых грубых априорных знаний о поверхности модели головы. Такой подход используется практически во всех современных методах создания аватара как головы, так и всего тела. Данный этап используется в разработанном методе и подробно описан в разделе 2 диссертации.

Раздел завершается выводами, в которых отмечаются преимущества и недостатки существующих решений. Формируется набор требований к разрабатываемому методу создания

цифрового аватара головы человека. Также приводится схематичное представление разработанного метода создания аватара головы, ключевые компоненты которого определяют структуру диссертации.

**Во втором разделе** диссертационного исследования представлен подробный обзор параметрической модели головы FLAME и описан классический алгоритм оценки параметров модели по RGB изображению, который применяется во многих современных работах по созданию цифрового аватара головы на этапе предобработки. Также представлен разработанный алгоритм оценки параметров параметрической модели FLAME по RGBD изображению.

Здесь и далее под параметрической моделью головы понимается некоторое представление единого формата (например, полигональная сетка с единой топологией), которое может контролироваться с помощью набора независимых параметров.

*В первом подразделе второго раздела* диссертационного исследования обосновывается необходимость использования параметрической модели головы для методов создания цифрового аватара. Необходимость связана с тем, что с помощью параметрической модели можно получить «грубую» априорную информацию, описывающую поверхность головы. Такую информацию затем можно использовать для обуславливания метода создания параметризованного аватара.

*Во втором подразделе второго раздела* диссертационного исследования приведено подробное описание взятой за основу параметрической модели головы FLAME. Модель FLAME представлена в виде полигональной сетки, которая контролируется параметрами, описывающими форму головы, выражение лица и позу. Под параметрами понимаются вещественнозначные вектора фиксированного размера. Так, модель FLAME описывается функцией  $M(\vec{\beta}, \vec{\theta}, \vec{\psi}): R^{|\vec{\beta}| \times |\vec{\theta}| \times |\vec{\psi}|} \rightarrow R^{3N}$ , которая принимает на вход параметры, описывающие форму  $\vec{\beta} \in R^{|\vec{\beta}|}$ , позу  $\vec{\theta} \in R^{|\vec{\theta}|}$ , выражение лица  $\vec{\psi} \in R^{|\vec{\psi}|}$ , и возвращает  $N$  вершин:

$$M(\vec{\beta}, \vec{\theta}, \vec{\psi}) = W(T_p(\vec{\beta}, \vec{\theta}, \vec{\psi}), J(\vec{\beta}; J, \bar{T}, S), \vec{\theta}, \mathcal{W}), \quad (1)$$

где  $M(\vec{\beta}, \vec{\theta}, \vec{\psi})$  – результирующая полигональная сетка для вектора формы  $\vec{\beta}$ , вектора позы  $\vec{\theta}$ , вектора выражения лица  $\vec{\psi}$ ,

$T_p(\vec{\beta}, \vec{\theta}, \vec{\psi})$  – шаблонная полигональная сетка с добавлением смещений в соответствии с формой, позой и выражением лица,

$J$  – положения подвижных сочленений (шея, челюсть, глаза) для данного субъекта,

$W$  – функция LBS, применяется для поворота вершин полигональной сетки  $T_p$  вокруг суставов  $J$ , выполняя линейное сглаживание весами смешивания  $\mathcal{W} \in R^{K \times N}$ .

Модель FLAME не включает в себя текстурную составляющую в отличие от популярной модели Basel Face Model (BFM), однако в 2020 году на платформе GitHub был представлен проект, в котором на основе текстурного пространства BFM и набора данных FFHQ (Flickr-Faces-HQ Dataset) было сформировано текстурное пространство модели FLAME с использованием алгоритма PCA.

*В третьем подразделе второго раздела* приведено описание классического алгоритма оценки параметров модели FLAME по RGB изображению. Для реализации процедуры производится вычисление бинарной маски, отделяющей голову от фона, и ключевых точек лица. Эти данные фиксируются и принимаются за истинную разметку. Вследствие чего итоговый результат ограничен и напрямую зависит от качества работы используемых сторонних алгоритмов. Затем выполняется инициализация нейтральными значениями параметров формы головы, выражения лица и позы модели FLAME, внутренних и внешних параметров камеры, параметров текстуры и освещения, которые подлежат оптимизации путем применения алгоритма итеративного градиентного спуска. Функция потерь (англ. loss function) для процедуры оптимизации имеет следующий вид:

$$\begin{aligned}
LOSS_{total}^{RGB} = & w_1 \cdot \frac{|img_{predict} - img_{GT}| \odot mask_{GT}}{\sum_{i=0}^{N-1} \sum_{j=0}^{M-1} mask_{GT_{i,j}}} + w_2 \cdot \frac{1}{2 \cdot L} \sum_{n=0}^{L-1} \sum_{m=0}^1 |lmk_{2D_{predict_{n,m}}} - \\
& lmk_{2D_{GT_{n,m}}}| + w_3 \cdot \sum_{k=0}^{|\vec{\beta}|} \left( \frac{\beta_k}{\sigma_{shape}} \right)^2 + w_4 \cdot \sum_{p=0}^{|\vec{\psi}|} \left( \frac{\psi_p}{\sigma_{expr}} \right)^2 + w_5 \cdot \sum_{q=0}^{|\vec{\tau}|} \left( \frac{\tau_q}{\sigma_{tex}} \right)^2 + w_6 \cdot \\
& \sum_{k=0}^{9K-1} \left( R(\vec{\theta}_k) - R(\vec{\theta}_k^*) \right)^2 + w_7 \cdot \sum_{l=0}^2 (\theta_{eye_l} - \theta_{eye_{l+3}})^2 \rightarrow \min_{\vec{\theta}, \vec{\beta}, \vec{\psi}, \vec{\tau}, \vec{\lambda}, \vec{cam}_{int}, \vec{cam}_{ext}}, \quad (2)
\end{aligned}$$

где  $img_{predict}$  – изображение, полученное после процедуры рендеринга с использованием текстуры,

$img_{GT}$  – исходное изображение,

$mask_{GT}$  – бинарная маска для исходного изображения,

$lmk_{2D_{predict}}$  – 2D координаты ключевых точек, спроецированных на отрендеренное изображение,

$lmk_{2D_{GT}}$  – истинные 2D координаты ключевых точек,

$\vec{\beta}, \vec{\psi}, \vec{\tau}$  – вектора коэффициентов формы, выражения лица и текстуры,

$\vec{\theta}$  – вектор параметров позы ( $\theta_{eye}$  – параметры позы для глаз;  $\vec{\theta}^*$  – параметры позы покоя,

$K = 4$  – количество суставов),

$\vec{\lambda}$  – вектор параметров освещения,

$\vec{cam}_{int}, \vec{cam}_{ext}$  – вектора внутренних и внешних параметров камеры,

$\sigma_{shape}, \sigma_{expr}, \sigma_{tex}$  – значения дисперсии для параметров векторов формы, выражения лица и текстуры,

$R(\vec{\theta})$  – функция, отображающая вектор позы  $\vec{\theta}$  в вектор, содержащий конкатенированные элементы всех соответствующих матриц вращения. Так,  $R_k(\vec{\theta})$  – это  $k$ -ый элемент  $R(\vec{\theta})$ ,

$w_1, w_2, w_3, w_4, w_5, w_6, w_7$  – взвешивающие коэффициенты для каждого слагаемого.

Полученную в ходе оценки параметров модель головы можно анимировать путем замены параметров выражения лица и/или параметров позы на требуемые и объединения полученного набора изображений в видеопоследовательность. Параметры позы и выражения лица можно получить также с помощью алгоритма оценки параметров, либо сгенерировать в допустимом диапазоне.

В четвертом подразделе второго раздела приведено описание разработанного алгоритма оценки параметров модели FLAME по RGBD изображению [1]. Ключевым преимуществом предложенного алгоритма является использование 3D разметки в виде облака точек, захваченного с помощью стереокамеры. На рисунке 2 схематично представлен предложенный алгоритм оценки параметров и анимации параметрической модели FLAME по RGBD изображению.

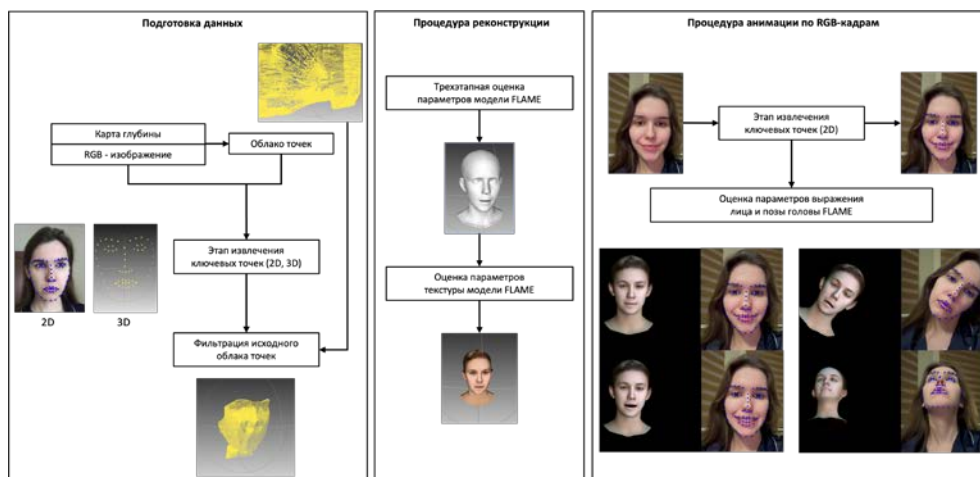


Рисунок 2 – Предложенный алгоритм оценки параметров и анимации параметрической модели FLAME по RGBD изображению



После выполнения подготовки входных данных производится оценка параметров модели FLAME и камеры. Процедура разбита на три блока. Первый блок отвечает за оценку параметров позы  $\vec{\theta}$  и камеры путем минимизации функции потерь, которая вычисляется по формуле (3). Второй блок отвечает за оценку параметров формы  $\vec{\beta}$  и выражения лица  $\vec{\psi}$ , минимизируя функцию потерь, вычисляемую по формуле (4). И, наконец, третий блок отвечает за оценку параметров формы  $\vec{\beta}$ , минимизируя функцию потерь, вычисляемую по формуле (5). Такое разделение на блоки позволяет сначала выполнить правильное позиционирование и лишь затем уточнять детали модели головы.

$$Loss_1^{RGBD} = \sqrt{\frac{1}{3 \cdot N} \sum_{i=0}^{N-1} \sum_{j=0}^2 \left( lmk_{3D_{GT_{i,j}}} - lmk_{3D_{predict_{i,j}}} \right)^2} \rightarrow \min_{\vec{\theta}, \vec{cam}_{int}, \vec{cam}_{ext}}, \quad (3)$$

где  $lmk_{3D_{GT}}$  – истинные 3D координаты ключевых точек,

$lmk_{3D_{predict}}$  – 3D координаты ключевых точек полигональной сетки модели FLAME,

$$Loss_2^{RGBD} = w_1 \cdot \sqrt{\frac{1}{3 \cdot N} \sum_{i=0}^{N-1} \sum_{j=0}^2 \left( lmk_{3D_{GT_{i,j}}} - lmk_{3D_{predict_{i,j}}} \right)^2} + w_2 \cdot \frac{1}{|PCD_{GT}|} \sum_{p_1 \in PCD_{GT}} \min_{p_2 \in M(\vec{\beta}, \vec{\theta}, \vec{\psi})} \|p_1 - p_2\|_2^2 \rightarrow \min_{\vec{\beta}, \vec{\psi}, \vec{\theta}, \vec{cam}_{int}, \vec{cam}_{ext}}, \quad (4)$$

где  $PCD_{GT}$  – отфильтрованное облако точек, которое представлено неупорядоченным набором 3D точек,

$|\cdot|$  – мощность множества,

$M(\vec{\beta}, \vec{\theta}, \vec{\psi})$  – полигональная сетка параметрической модели FLAME,

$w_1, w_2$  – взвешивающие коэффициенты для каждого слагаемого.

$$Loss_3^{RGBD} = \frac{1}{|PCD_{GT}|} \sum_{p_1 \in PCD_{GT}} \min_{p_2 \in M(\vec{\beta}, \vec{\theta}, \vec{\psi})} \|p_1 - p_2\|_2^2 \rightarrow \min_{\vec{\beta}}, \quad (5)$$

Затем производится оценка параметров текстуры  $\vec{\tau}$  и освещения  $\vec{\lambda}$  путем минимизации функции потерь, которая вычисляется по формуле (6).

$$Loss_4^{RGBD} = \frac{|img_{predict} - img_{GT}| \odot mask_{GT}}{\sum_{i=0}^{N-1} \sum_{j=0}^2 mask_{GT_{i,j}}} \rightarrow \min_{\vec{\tau}, \vec{\lambda}}, \quad (6)$$

Для набора данных, участвующего в экспериментальном исследовании алгоритма оценки параметров параметрической модели по RGBD изображению, было получено среднее значение функции потерь  $Loss_3$  равное  $4 \cdot 10^{-8}$ . Так как функция потерь  $Loss_3$  соответствует последнему этапу трехэтапной процедуры оценки параметров параметрической модели FLAME, то ее значение можно рассматривать как меру качества. Ее физический смысл состоит в измерении среднего значения квадрата расстояния между каждой точкой из исходного облака точек, полученного с использованием стереокамеры, и ближайшей к ней гранью полигональной сетки, полученной в ходе оптимизации. Так как захваченные облака точек представлены в реальном масштабе, а их значения измеряются в метрах, то  $\sqrt{Loss_3^{RGBD}} = 0,0002$  м. Далее будем называть эту величину точностью трехмерной реконструкции.

На рисунке 3 представлены слева-направо: результат работы алгоритма оценки параметров модели по RGB изображению, результат работы метода оценки параметров DECA, который используется в некоторых актуальных методах создания аватара головы и описан в разделе 1 диссертации, и результат работы предложенного алгоритма оценки параметров модели по RGBD кадру. Для оценки параметров

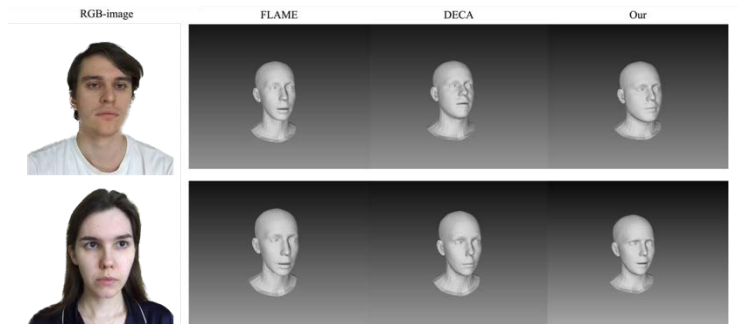


Рисунок 3 – Сравнение результатов реконструкции

методом DECA использовалось RGB изображение и 2D ключевые точки.

Раздел завершается выводами, где отмечено, что предложенный алгоритм оценки параметров модели по RGBD снимку дает более качественные результаты, которые близки к истинной форме головы, в отличие от метода DECA и классического алгоритма оценки параметров модели FLAME по RGB снимку. При этом предложенный алгоритм достигает высокой точности трехмерной реконструкции равной 0,2 мм. Однако применение такого подхода ограничено наличием стереокамеры, которая способна формировать облака точек с высокой точностью.

**В третьем разделе** диссертационного исследования приведено описание разработанной параметрической модели головы человека на основе нейросетевой модели представления поверхности CNeRF (Conditional Neural Radiance Fields) и двумерного нейронного рендеринга [4]. Описан модуль генерации синтетических данных в реальном времени. Представлены порядок проведения и результаты экспериментальных исследований.

Параметрическая модель может быть полезна в качестве одного из этапов метода создания цифрового аватара головы человека, поскольку может использовать те же процедуры оценки параметров, что и оригинальная модель FLAME, при этом являясь богатым источником априорной информации в случае применения методик переноса обучения (англ. transfer learning).

*В первом подразделе третьего раздела* приведено описание архитектуры разработанной параметрической модели головы, включая описание её компонент: нейросетевой модели представления поверхности CNeRF, алгоритма двумерного нейронного рендеринга, а также используемого способа обучения.

Архитектура состоит из нейронных полей излучения, обусловленных параметрами параметрической модели головы FLAME, и нейронной сети, которая выполняет двумерный нейронный рендеринг. Условные нейронные поля излучения создают низкоразмерную карту признаков с использованием алгоритма объеметрического рендеринга. Нейронная сеть, выполняющая двумерный нейронный рендеринг, из полученной карты признаков формирует итоговое изображение. Такой подход в отличие от ранних методов, основанных на нейронной модели NeRF, позволяет ускорить процедуру рендеринга, так как карта признаков имеет небольшое разрешение в отличие от итогового изображения. За счет обуславливания многослойного перцептрона, в модели NeRF становится возможным контроль и варьирование параметров формы, выражения лица и позы головы, что отсутствует в базовой версии NeRF, предназначенной для представления статической сцены. На рисунке 4 приведено схематичное представление архитектуры предложенной параметрической модели.

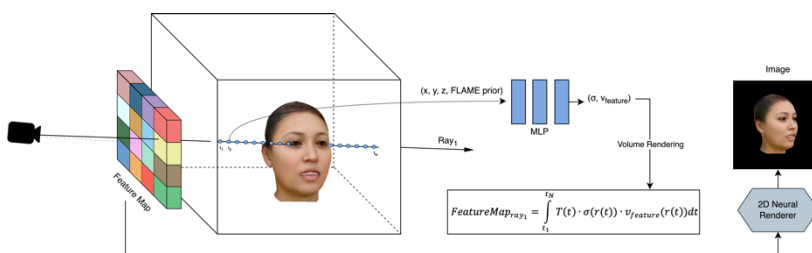


Рисунок 4 – Схематичное представление предложенной архитектуры параметрической модели головы

ускорить процедуру рендеринга, так как карта признаков имеет небольшое разрешение в отличие от итогового изображения. За счет обуславливания многослойного перцептрона, в модели NeRF становится возможным контроль и варьирование параметров формы, выражения лица и позы головы, что отсутствует в базовой версии NeRF, предназначенной для представления статической сцены. На рисунке 4 приведено схематичное представление архитектуры предложенной параметрической модели.

На рисунке 5 представлена архитектура нейронной сети, которая выполняет двумерный нейронный рендеринг. Основным отличием, которое позволило ускорить процесс сходимости и улучшить качество рендеринга в отличие от подхода, представленного Hong Y. с соавторами в своей работе, является блок повышения пространственной дискретизации (увеличивает пространственное разрешение входной карты признаков и уменьшает количество каналов). Этот блок разделен на две части, которые выполняют задачи выделения локального и глобального контекстов, для этого:

1. Входная карта признаков проходит через два отдельных блока Squeeze-and-Excitation, которые присваивают вес каждому каналу входного тензора, позволяя тем самым модели сосредоточиться на наиболее значимых каналах карты признаков для

решения поставленной задачи (под задачами в данном случае понимается выделение локального и глобального контекста). Полученные карты признаков проходят через последовательность свёрточных слоёв в связке с функцией активации LeakyReLU.

2. В первой части блока для выделения глобального контекста используются расширенные свёрточные слои (англ. dilated convolution layers) с размером ядра  $3 \times 3$ . За счёт большого рецептивного поля (англ. receptive fields) такие слои позволяют концентрироваться на глобальных характеристиках.
3. Во второй части блока для выделения локального контекста используются свёрточные слои с размером ядра  $1 \times 1$ , что позволяет фокусироваться на более мелких деталях.
4. Полученные из двух блоков карты признаков затем суммируются и поступают на вход необучаемого слоя Pixel Shuffle, который уменьшает глубину входной карты признаков и увеличивает разрешение за счет перестановок ее значений.
5. Полученная на (4) этапе карта признаков обрабатывается сглаживающим фильтром.

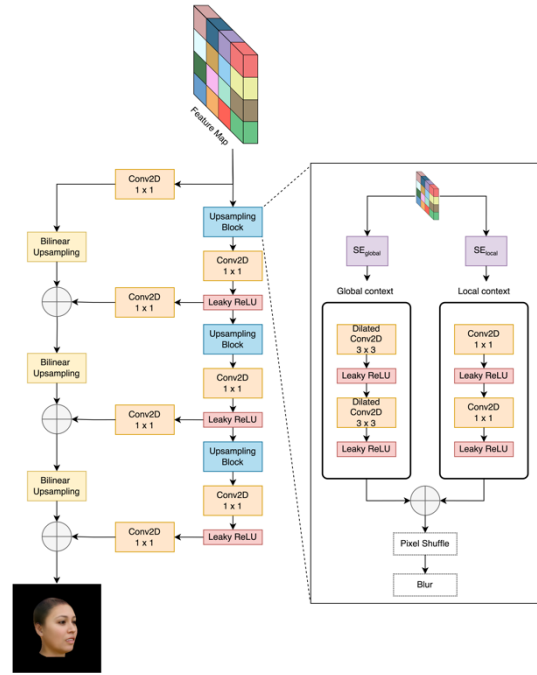


Рисунок 5 – Архитектура нейронной сети, которая выполняет двумерный нейронный рендеринг

На рисунке 6 представлен пример карт признаков после свёрточных слоёв в блоке повышения дискретизации, демонстрирующий работу частей блока для выделения глобального и локального контекста, и подтверждающий их назначение.

Оптимизация параметров предложенной параметрической модели головы производится с использованием итеративного алгоритма градиентного спуска путем минимизации функции потерь, которая вычисляется по формуле (7).

$$Loss_{total} = w_1 \cdot \frac{\| (img_{GT} - img_{predict}) \odot img_{mask} \|^2}{\sum_{i=0}^{N-1} \sum_{j=0}^{M-1} img_{mask}} + w_2 \cdot \sum_{j=0}^{L-1} \|\phi_j(img_{GT}) - \phi_j(img_{predict})\|^2 + \sum_{k \in \{\vec{\beta}, \vec{\theta}, \vec{\psi}, \vec{\tau}, \vec{\lambda}\}} w_k \cdot \left\| \overrightarrow{v_k^{offset}} \right\|^2 \rightarrow$$

$$\min_{\substack{CNeRF \text{ parameters,} \\ 2D \text{ Neural Renderer parameters,} \\ \overrightarrow{v_k^{offset}}, k \in \{\vec{\beta}, \vec{\theta}, \vec{\psi}, \vec{\tau}, \vec{\lambda}\}}} , \quad (7)$$

где  $img_{GT}$  – исходное изображение,

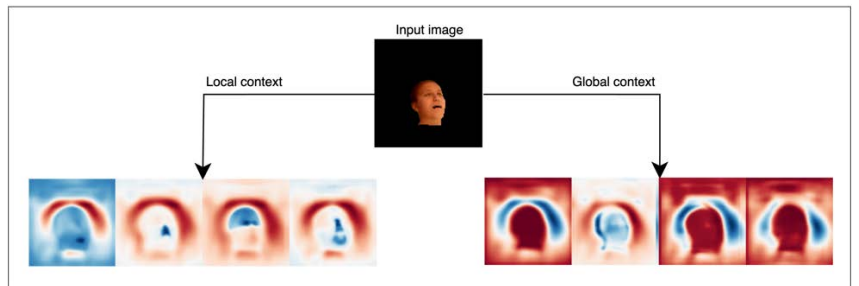


Рисунок 6 – Примеры карт признаков после свёрточных слоёв последнего уровня для частей, выделяющих глобальный и локальный контекст.

$img_{predict}$  – синтезированное изображение,

$img_{mask}$  – маска сегментации головы исходного изображения,

$\phi_j(*)$  – активация  $i$ -го слоя предобученной нейронной сети архитектуры VGG16,

$\vec{v}_{\vec{\beta}}^{offset}, \vec{v}_{\vec{\theta}}^{offset}, \vec{v}_{\vec{\psi}}^{offset}, \vec{v}_{\vec{\tau}}^{offset}, \vec{v}_{\vec{\lambda}}^{offset}$  – вектора, описывающие смещение от

исходных векторов коэффициентов формы, параметров позы (шея, челюсть, глазные яблоки), коэффициентов выражения лица, текстурных коэффициентов и параметров освещения,

$w_1; w_2; w_k, k \in \{\vec{\beta}, \vec{\theta}, \vec{\psi}, \vec{\tau}, \vec{\lambda}\}$  – взвешивающие коэффициенты.

Во втором подразделе третьего раздела описывается процесс формирования синтетического набора данных, основанного на открытом наборе данных лиц FFHQ, и параметрической модели головы FLAME. Для каждого изображения из набора вычисляются ключевые точки лица и бинарная маска сегментации (включая волосы), после чего выполняется процедура оценки параметров модели FLAME, описанная в подразделе 2.3 диссертации. На рисунке 7 представлен результат оценки параметров модели FLAME для нескольких изображений из набора.

Во время обучения параметров модели к векторам FLAME-коэффициентов формы, выражения лица и текстуры каждого изображения добавляется шум, сгенерированный по равномерному распределению в заранее установленном диапазоне значений векторов для получения реалистичных параметров. Параметры, отвечающие за позу (поворот шеи, смещение челюсти), генерируются из равномерного распределения в допустимых диапазонах. Внешние параметры камеры также генерируются из равномерного распределения в рамках указанного диапазона.

В третьем подразделе третьего раздела описана стратегия обучения разработанной параметрической модели головы человека. На рисунке 8 представлена схема итерации процесса обучения параметрической модели головы, в ходе которого производится оптимизация всех обучаемых параметров с использованием синтетического набора данных, генерируемого в реальном времени с использованием параметрической модели головы FLAME.

Для проведения экспериментальных исследований обучения параметрической модели головы была предложена стратегия обучения, при которой вероятность включения синтетических данных рассчитывается в зависимости от номера итерации согласно формуле (8). Была выдвинута гипотеза, что такой подход к обучению позволит добавить в обучаемую параметрическую

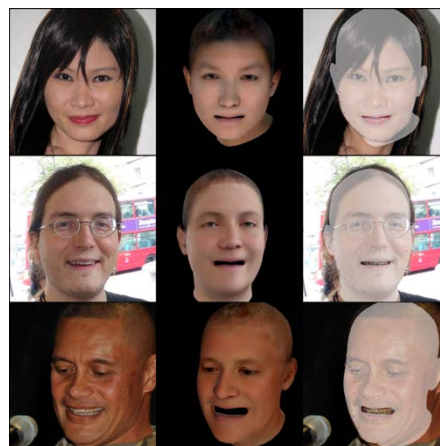


Рисунок 7 – Результат оценки параметров модели FLAME (столбцы 2, 3) и оригинальное изображение (столбец 1)

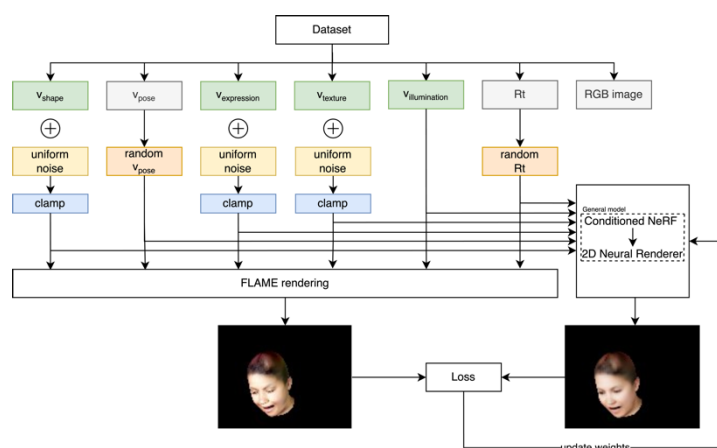


Рисунок 8 – Схематическое представление итерации процесса обучения параметрической модели головы человека с использованием синтетического набора данных

модель априорную информацию о высокоуровневых признаках лиц людей, отсутствующих в изображениях, полученных в результате рендеринга модели FLAME.

$$p = \begin{cases} 1, & \text{если } n \leq M, \\ \left(1 - \frac{n-M}{N-M}\right) \cdot (p_{\max} - p_{\min}) + p_{\min}, & \text{если } n > M. \end{cases} \quad (8)$$

где  $n$  – номер текущей итерации,  
 $M$  – количество итераций для обучения исключительно на синтетических данных,  
 $N$  – общее количество итераций обучения,  
 $p_{\max}$  – максимальное значение вероятности включения синтетических данных,  
 $p_{\min}$  – минимальное значение вероятности включения синтетических данных,  
 $p$  – вероятность использования синтетических данных в процессе обучения,  
 $1 - p$  – вероятность использования данных из FFHQ в процессе обучения.

На рисунке 9 представлена схема процесса обучения параметрической модели головы, в ходе которого производится оптимизация всех обучаемых параметров метода на наборе данных FFHQ, который предварительно проходит процедуру оценки параметров модели FLAME.

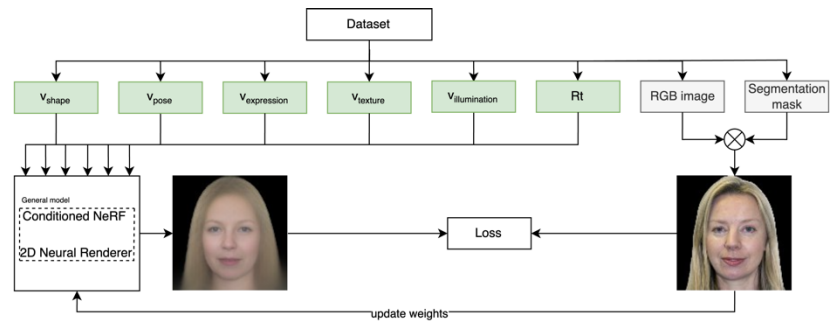


Рисунок 9 – Схематическое представление итерации процесса обучения параметрической модели головы человека с использованием набора данных FFHQ

В четвертом

подразделе третьего раздела были описаны порядок проведения и результаты экспериментальных исследований. Исходя из представленных результатов, проведенных экспериментальных исследований, были сделаны следующие выводы:

1. Разработанная модель обладает способностью синтезировать изображения, соответствующие изображениям, синтезированным при помощи модели FLAME, что подтверждают представленные показатели качества ( $PSNR \approx 30$  дБ) и полученные изображения-проекции для тестовых данных.
2. Разработанный блок в архитектуре нейронной сети, выполняющей двумерный нейронный рендеринг, позволил уменьшить количество итераций для обучения модели более чем в 2 раза по сравнению с методом, представленным Hong Y. с соавторами в своей работе.
3. Использование модифицированной текстуры приводит к тому, что изображения, синтезируемые при помощи разработанной модели, имеют большее сходство с реальной человеческой головой.
4. Использование набора данных FFHQ в процессе обучения приводит к тому, что результаты синтеза обладают большей фотореалистичностью.
5. Применение аугментации, моделирующей изменение расстояния по оси  $z$  головы от камеры, к изображениям из набора данных FFHQ приводит к уменьшению эффекта переобучения на результат обучения разработанной модели.
6. Использование стратегии обучения разработанной модели с  $p_{\min} = p_{\max} = 0$  приводит к худшим результатам, чем обучение с  $p_{\min} = p_{\max} = 0,25$ . Использование стратегии с параметрами  $p_{\min} = p_{\max} = 0,25$  позволяет снизить эффект переобучения модели.
7. Использование стратегии обучения разработанной модели с  $p_{\min} = p_{\max} = 0,25$  приводит к худшим результатам, чем обучение с  $p_{\max} = 0,25$  и  $p_{\min} = 0$ . Использование стратегии с параметрами  $p_{\min} = 0$ ,  $p_{\max} = 0,25$  позволяет снизить эффект переобучения до минимума, так как при  $p_{\min} = p_{\max}$  параметрическая модель обучается отделять параметры, соответствующие синтетическим и реальным

данным. Качественно обученная модель должна обладать способностью синтезировать изображения с высокой консистентностью, то есть при изменении параметров, отвечающих за наклон или поворот головы, должен изменяться только поворот или наклон головы на синтезируемом изображении, и это изменение параметра не должно затрагивать другие свойства человеческого лица (форма, выражение лица).

#### В четвертом разделе

диссертационного исследования приведено описание разработанного метода создания параметризованного аватара головы человека [4]. Описываются порядок проведения и результаты экспериментальных исследований метода. В частности, описывается сформированный набор настроек метода, предлагаемые варианты аугментации данных, что в совокупности позволяет разработанному методу решать поставленную задачу эффективнее. Представлен полный цикл создания и анимации аватара головы, что может быть полезно в прикладных сценариях использования разработанного метода. А также приведено сравнение предложенного метода с другими современными и актуальными методами, позиционируемыми авторами, как state-of-the-art решения.

В первом подразделе четвертого раздела приводится описание разработанного метода создания параметризованного аватара головы человека. Так, разработанный метод использует архитектуру предложенной параметрической модели головы, подробно описанную в разделе 3 диссертации, и её веса для инициализации параметров. Такой способ инициализации позволяет использовать в том числе короткую монокулярную видеопоследовательность (порядка 1000 кадров) и небольшое число итераций стохастического градиентного спуска, чтобы получить возможность синтезировать новые согласованные виды и мимику для целевого лица с минимальным значением ошибки. На рисунке 10 представлена схема итерации процесса обучения предлагаемого метода создания параметризованного аватара головы человека. На рисунке 11 представлена схема полного цикла создания аватара головы человека при помощи разработанного метода.

Во втором подразделе четвертого раздела приведены возможные приложения разработанного метода создания параметризованного аватара, а именно – синтез новых видов и перенос выражения лица. Важно отметить, что процедуру переноса выражения лица можно выполнить по видеопоследовательности, содержащей мимику другого человека.

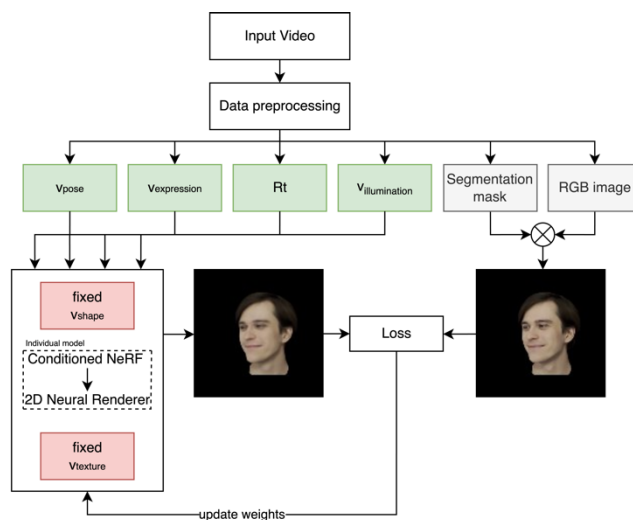


Рисунок 10 – Схематическое представление итерации процесса обучения предлагаемого метода создания параметризованного аватара головы человека

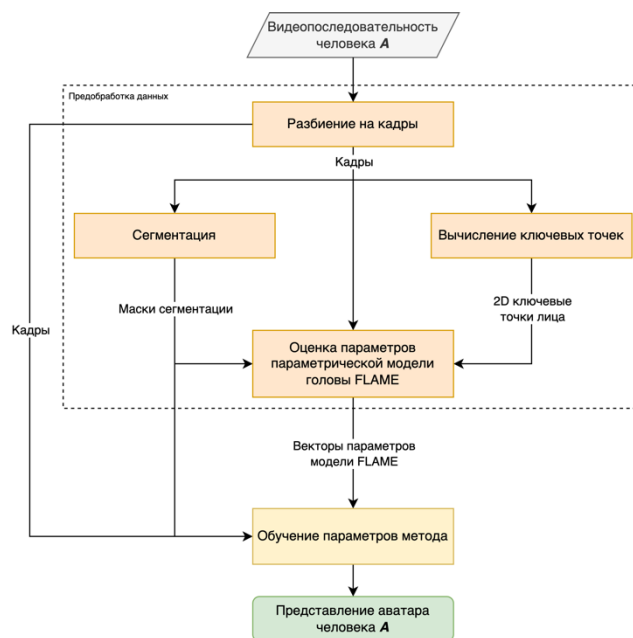


Рисунок 11 – Схема полного цикла создания аватара головы человека при помощи разработанного метода

Для выполнения процедуры синтеза новых видов, под которой понимается генерация изображений аватара с различных точек обзора, необходимо выполнить замену значений внешних параметров камеры (матрица поворота и вектор сдвига) на требуемые. На рисунке 12 схематично представлена процедура синтеза новых видов аватара.

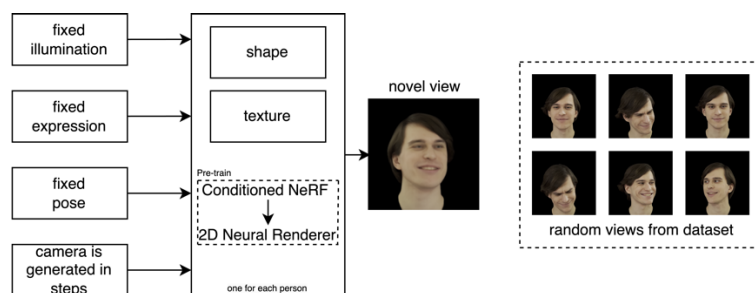


Рисунок 12 – Схематичное представление отдельной итерации синтеза новых видов аватара

Для выполнения процедуры переноса выражения лица с кадров видеопоследовательности необходимо: выполнить оценку параметров параметрической модели головы FLAME для всех кадров видеопоследовательности; выполнить замену вектора выражения лица на требуемый; опционально: выполнить замену параметров камеры и/или позы и/или освещения. На рисунке 13 приведено схематичное представление полного цикла процедуры анимации по видеопоследовательности (в т. ч. перенос выражения лица).

В третьем подразделе четвертого раздела предложен способ расширения исходного набора данных для обучения персонализированного параметризованного аватара головы человека. Так, предлагается использовать метод XVFI, предложенный Sim H. и соавторами, для интерполяции промежуточных кадров.

Для определения наиболее эффективного метода интерполяции динамических сцен было произведено сравнение существующих актуальных решений [2]. По результатам сравнения, было выявлено, что метод XVFI показывает наилучшее качество интерполяции для видеопоследовательностей с высокой вариативностью.

На основе полученных результатов было произведено исследование влияния интерполяции промежуточных кадров на результат трехмерной реконструкции, в том числе с использованием метода NeRF [3]. Согласно результатам исследований, расширение набора данных с использованием метода XVFI приводит к улучшению качества трехмерной реконструкции.

В четвертом подразделе четвертого раздела описаны порядок проведения и результаты экспериментальных исследований разработанного метода создания параметризованного аватара головы человека. В рамках исследований используется набор данных, содержащий 10 видеопоследовательностей. Первые 6 видеопоследовательностей опубликованы Zielonka W. и соавторами и содержат порядка 3000 кадров каждая. Вторые 2 видеопоследовательности опубликованы Grassal P.-W. и соавторами и содержат порядка 1000 кадров каждая. Последние 2 видеопоследовательности записаны самостоятельно и содержат порядка 800 кадров

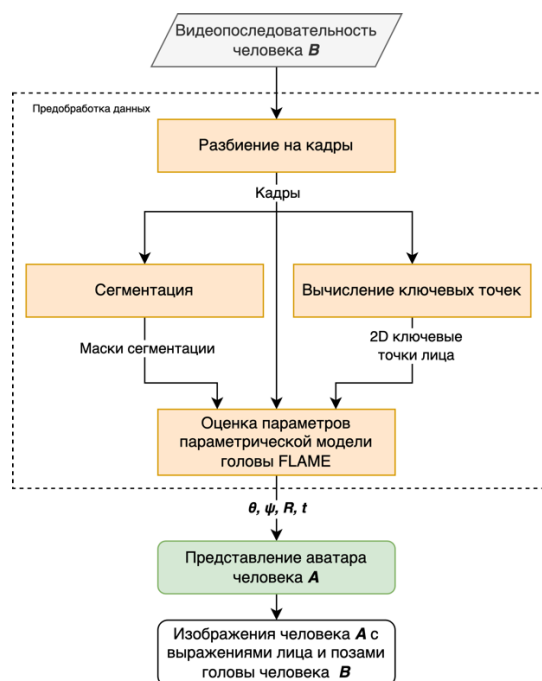


Рисунок 13 – Схема процедуры анимации аватара головы человека по входной видеопоследовательности

каждая.

Выбранные открытые наборы данных наиболее часто используются для оценки качества новых методов создания параметризованного аватара головы исследователями. Поэтому их использование в рамках экспериментальных исследований позволяет объективно оценить разработанный метод. В связи с тем, что такие наборы состоят из RGB-

видеопоследовательностей, для оценки параметров модели FLAME используется алгоритм оценки по RGB изображению.

Подраздел подробно описывает обширные экспериментальные исследования предложенного метода, в ходе которых было получено 170 аватаров. На основе проведенных исследований разработанного метода создания параметризованного аватара головы человека формируется итоговый набор настроек обучения:

1. Инициализация параметров: подобрана наилучшая конфигурация обучения разработанной параметрической модели головы человека, обеспечивающая высокое качество синтеза с использованием разработанного метода создания параметризованного аватара головы.
2. Аугментация: симуляция изменения расстояния от камеры до головы по оси  $z$  для изображений из обучающей выборки в процессе оптимизации параметров для аватара.
3. Использование стратегии, при которой вероятность включения синтетических данных в обучающую выборку линейно убывает с ростом количества итераций при оптимизации параметров для аватара.

Затем были проведены экспериментальные исследования для определения целесообразности использования предложенного способа аугментации (расширения) набора данных. Для каждой исходной видеопоследовательности в наборе данных был удален каждый второй кадр и затем восстановлен при помощи метода XVFI. Полученные видеопоследовательности были использованы для создания аватаров. Для тестовых выборок, соответствующих полученным аватарам, были вычислены показатели качества PSNR, SSIM и LPIPS, значения которых подтверждают целесообразность применения предложенного способа расширения набора данных. Так, например, среднее значение показателя качества PSNR по отношению к тестовой выборке увеличивается на 0,17 дБ.

В пятом подразделе четвертого раздела приведено сравнение разработанного метода создания аватара с актуальными существующими решениями, позиционируемыми авторами как state-of-the-art в исследуемой области, а именно Neural Head Avatar, IMAvatar, PointAvatar, INSTA.

Количественные результаты проведенного исследования свидетельствуют о том, что предложенный метод превосходит существующие по показателям качества PSNR и LPIPS на выбранном наборе данных. Разработанный метод позволяет синтезировать изображения-проекции со средним значением PSNR 26,29 дБ по отношению к изображениям из тестовой выборки. Зависимость качества синтеза изображений-проекций аватаров для тестовых выборок от времени, затраченного на обучение, для всех методов представлена на рисунке 14. Зависимость качества синтеза изображений-проекций аватаров для тестовых выборок от среднего времени, затраченного на генерацию одного кадра, для всех методов представлена на рисунке 15.

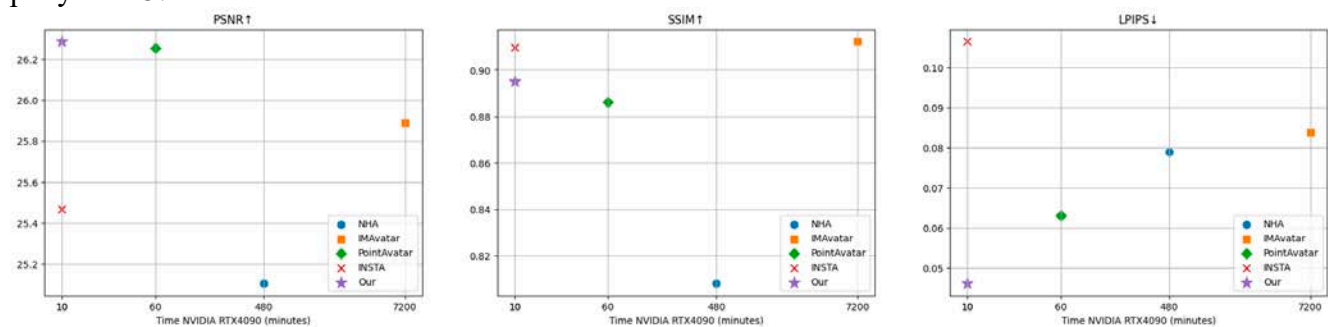


Рисунок 14 – Зависимость качества синтеза изображений-проекций аватаров для тестовых выборок от времени, затраченного на обучение, для всех методов



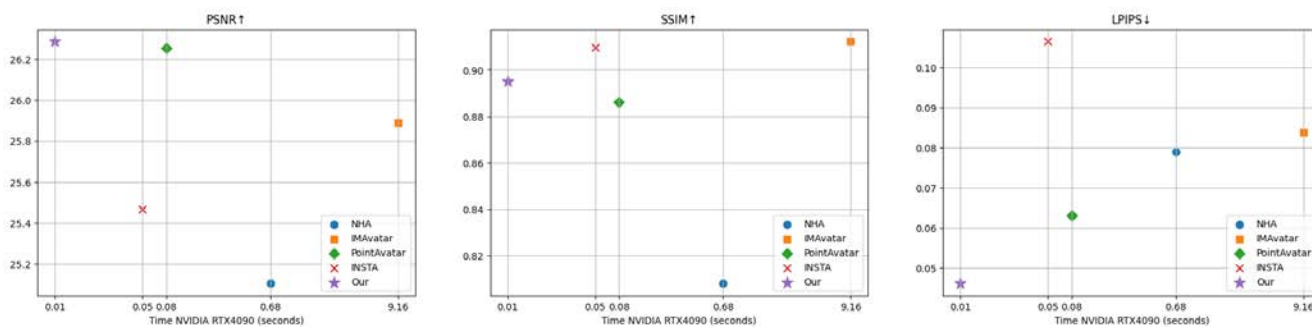


Рисунок 15 – Зависимость качества синтеза изображений-проекций аватаров для тестовых выборок от среднего времени, затраченного на генерацию одного кадра, для всех методов

На рисунке 16 слева представлен результат синтеза изображений-проекций для экземпляров данных из тестовых выборок. На рисунке 16 справа представлен результат процедуры переноса выражения лица. Визуальный анализ синтезированных кадров на рисунке 16 подтверждает высокое качество работы метода. Разработанный метод превосходит в среднем рассмотренные методы по значению PSNR на 0,6 дБ; превосходит лучший метод в сравнении по времени, необходимому на синтез одного изображения-проекции, на 84%; находится на одном уровне с лучшим методом по времени, необходимому для синтеза представления аватара.

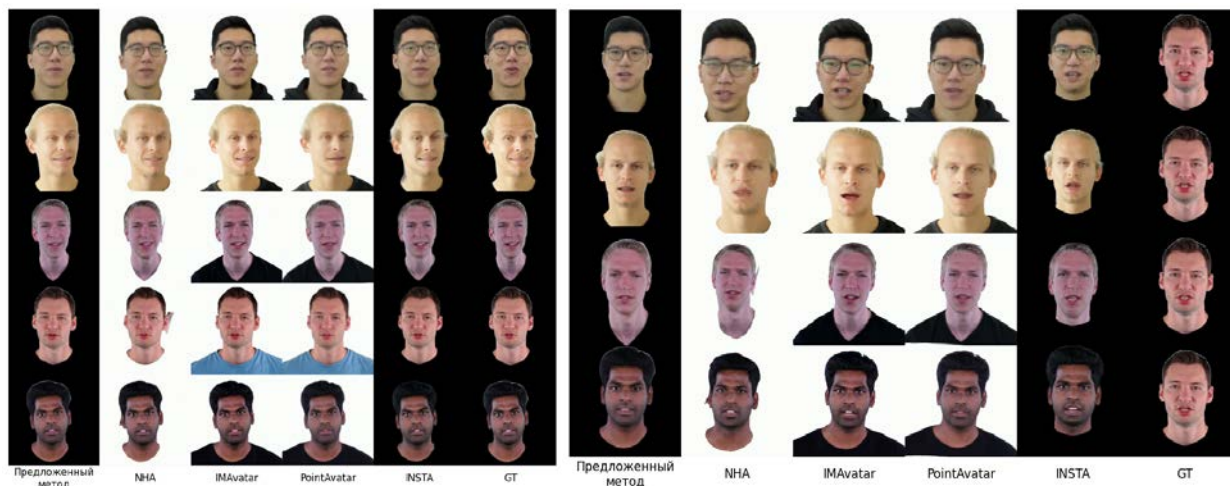


Рисунок 16 – Результат синтеза изображений-проекций для экземпляров данных из тестовых выборок (слева). Результат процедуры переноса выражения лица (справа)

## ЗАКЛЮЧЕНИЕ

В рамках диссертационной работы разработан и исследован новый метод создания параметризованного аватара головы человека на основе нейросетевой модели рендеринга, обеспечивающий при относительно низких вычислительных затратах процесса создания повышенное качество синтезируемых изображений-проекций. Основные результаты диссертационной работы:

1. Алгоритм оценки параметров параметрической модели головы FLAME с использованием RGBD изображения. Предложенный алгоритм достигает точности трехмерной реконструкции равной 0,2 мм.
2. Параметрическая модель головы человека на основе нейросетевой модели представления поверхности CNeRF, архитектуры сети двумерного нейронного рендеринга с блоком повышения пространственной дискретизации, позволяющим ускорить сходимость метода создания аватара (более чем в 2 раза), и синтетического набора данных, генерируемого в реальном времени. Разработанная модель позволяет синтезировать изображения-проекции со средним значением PSNR более 30 дБ по отношению к изображениям, полученным с использованием параметрической модели головы FLAME.

3. Метод создания аватара головы человека на основе разработанной параметризованной модели головы человека. Разработанный метод позволяет синтезировать изображения-проекции со средним значением PSNR 26,29 дБ по отношению к изображениям из тестовой выборки.
4. Способ аугментации (расширения) реального набора данных с изображениями головы человека с помощью интерполяции промежуточных кадров видеопоследовательности. Способ позволяет повысить качество (среднее значение показателя качества PSNR по отношению к тестовой выборке увеличивается на 0,17 дБ) синтезируемых изображений-проекций аватара для коротких видеопоследовательностей.
5. Результаты экспериментальных исследований разработанного метода создания параметризованного аватара, включающие сравнение разработанного метода с существующими актуальными (state-of-the-art) решениями. Разработанный метод превосходит в среднем рассмотренные методы по значению PSNR на 0,6 дБ; превосходит лучший метод из рассмотренных по времени, необходимому на синтез одного изображения-проекции, на 84%; находится на одном уровне с лучшим методом из рассмотренных по затратам времени, необходимым для создания аватара.

#### **Работы, опубликованные автором по теме диссертации**

*В изданиях, рекомендованных ВАК при Минобрнауки России и/или индексируемых реферативными базами данных Web of Science / Scopus:*

1. Kozlova, Y.K. Head model reconstruction and animation method using color image with depth information / Y.K. Kozlova, V.V. Myasnikov // Computer Optics. — 2024. — Vol. 48. P. 118–122. — DOI: 10.18287/2412-6179-CO-1334.
2. Ganeeva, Y. Comparison of methods for reconstructing intermediate video frames with a dynamic scene / Y. Ganeeva // IEEE Xplore 2022 VIII International Conference on Information Technology and Nanotechnology — 2022. — P. 1–5. — DOI: 10.1109/ITNT55410.2022.9848739.
3. Ganeeva, Y. The impact of intermediate video frames reconstruction step on the result of 3D reconstruction of objects / Y. Ganeeva, V. Myasnikov // IEEE Xplore 2022 VIII International Conference on Information Technology and Nanotechnology — 2022. — P. 1–5. — DOI: 10.1109/ITNT55410.2022.9848697.
4. Kozlova, Yu.Kh. Method for Creating Animatable Avatars Using Neural Radiance Fields and Two-Dimensional Neural Rendering / Yu.Kh. Kozlova // IEEE Xplore 2024 X International Conference on Information Technology and Nanotechnology — 2024. — P. 1–8. — DOI: 10.1109/ITNT60778.2024.10582377.

*В других изданиях:*

5. Ганеева<sup>1</sup>, Ю.Х. Сравнение методов реконструкции промежуточных кадров видеопоследовательностей с динамической сценой / Ю.Х. Ганеева // Информационные технологии и нанотехнологии. Сборник трудов по материалам VIII Международной конференции и молодежной школы. Изд-во Самарского Университета — Самара, 2022. — С. 31472.
6. Ганеева, Ю.Х. Влияние реконструкции промежуточных кадров видеопоследовательности на результат 3D-реконструкции объектов / Ю.Х. Ганеева, В.В. Мясников // Информационные технологии и нанотехнологии. Сборник трудов по материалам VIII Международной конференции и молодежной школы. Изд-во Самарского Университета — Самара, 2022. — С. 31492.
7. Козлова, Ю.Х. Метод реконструкции и анимации модели головы с использованием RGBD-изображения / Ю.Х. Козлова // Информационные технологии и нанотехнологии. Сборник трудов по материалам IX Международной конференции и молодежной школы. Изд-во Самарского Университета — Самара, 2023. — С. 30612.
8. Козлова, Ю.Х. Метод создания анимируемых аватаров с использованием нейронных полей излучения и двухмерного нейронного рендеринга / Ю.Х. Козлова // Информационные технологии и нанотехнологии. Сборник трудов по материалам X Международной конференции и молодежной школы. Изд-во Самарского Университета — Самара, 2024.

---

<sup>1</sup> Фамилия Ганеева изменена на фамилию Козлова (свидетельство о заключении брака П-ЕР № 841954, выдано отделом Дворцом бракосочетания городского округа Самара управления ЗАГС Самарской области, 20.08.2022г.).