

ОТЗЫВ

официального оппонента доктора технических наук, доцента
Дементьева Виталия Евгеньевича на диссертацию Козловой Юлии Ханифовны на
тему «Метод создания параметризованного аватара головы человека на основе
нейросетевой модели рендеринга», представленную на соискание ученой степени
кандидата технических наук по специальности 2.3.8 – Информатика и информаци-
онные процессы

1. Актуальность темы диссертации работы

В последнее время широкое распространение получили разнообразные системы виртуальной, смешанной и дополненной реальности. Такие системы востребованы в различных образовательных и исследовательских проектах, включая интерактивное освоение новых инструментов и технологий, при удаленном взаимодействии сотрудников друг с другом, дистанционном управлении роботизированными и беспилотными платформами и т.п. Качество и востребованность подобных систем во многом обусловлены тем, насколько в них достоверно воссоздаются различные объекты реального мира. В связи с этим значительное число исследований во всем мире посвящено тематике создания и исследования так называемых «цифровых двойников» - объектов виртуальной реальности, имитирующих свойства предметов, механизмов и даже систем реального мира. Использование подобных цифровых моделей действительно во многих случаях позволяет существенно упростить изучение свойств реальных объектов, в том числе при их взаимодействии с окружающим миром.

Одним из ключевых объектов нашего мира, наиболее востребованных при имитации в виртуальной реальности, является сам человек. Создание правдоподобных цифровых моделей конкретных людей позволяет решать ряд важных и востребованных задач, связанных, например, с обеспечением телеприсутствия человека при дистанционном общении, образовательных проектах, медицинских консульта-



циях, иных видах интерактивного взаимодействия, в том числе в условиях ограниченной пропускной способности каналов связи. К сожалению, в настоящее время создание таких цифровых моделей требует значительных временных и финансовых затрат, поскольку сопряжено с использованием специализированного дорогостоящего оборудования.

В связи со сказанным диссертационная работа Козловой Ю.Х., посвященная решению задач создания параметризованного аватара человека, то есть его фотореалистичного изображения, управляемого значениями параметров мимики и позы, является актуальной.

2. Структура и объём диссертации

Диссертация Козловой Ю.Х. состоит из введения, четырех разделов, заключения, списка литературы из 102 наименований и одного приложения. Текст диссертации изложен на 209 страницах машинописного текста (включая 77 страниц приложения), в т. ч. 59 рисунков, 2 таблицы. Приложение содержит 83 изображения, отражающих результаты синтеза аватар разных людей при различных параметрах.

Во введении обоснована актуальность работы, сформулированы ее цель и задачи. Проведён краткий обзор известных методов и инструментальных средств построения цифровых моделей человеческой головы, сформулированы основные проблемы. Обосновывается соответствие диссертации паспорту научной специальности. Приводятся перечень научных результатов, научная новизна, теоретическая и практическая значимость, а также положения, выносимые на защиту.

В первой главе приведена краткая классификация методов создания цифрового аватара головы человека. В частности, описаны способы представления поверхности цифрового аватара, включая явные (облако точек, полигональная сетка, воксельная сетка) и неявные (SDF, OF, NeRF), варианты обобщенности итоговой модели (общая или персональная модель), требования к набору данных для обучения (3D набор данных, разноразмерные плоские RGB изображения, синтетические наборы данных), варианты управления параметрами модели (явные и неявные). Описаны наиболее значимые с точки зрения автора решения в области создания цифровых

аватаров головы человека. Сравнительный анализ этих решений проведен в таблице 1 на страницах 34–36. По результатам этого анализа делается вывод о необходимости использования априорной информации о поверхности головы, полученной в результате оценки параметров модели, в нейросетевой модели, неявно описывающей цифровой аватар головы человека. Для решения этой задачи предлагается использовать параметрическую модель головы на основе нейросетевой модели представления поверхности CNeRF, которая формирует общее представление и предоставляет стартовые значения весов для разработанного автором метода создания цифрового аватара головы человека.

Во второй главе подробно рассмотрена параметрическая модели головы FLAME, проанализированы параметры этой модели и подробно исследован классический алгоритм оценки параметров модели FLAME по отдельному RGB кадру. По результатам этого исследования делается вывод о недостатке этого алгоритма, связанном с отсутствием при оптимизации информации о 3D координатах ключевых точек, что в результате приводит к недостаточной схожести формы головы. Для преодоления этого недостатка предлагается оригинальный алгоритм оценки параметров модели FLAME с использованием RGBD изображения. В рамках этого алгоритма вводится и рассматривается трехэтапная процедура оценка параметров модели и камеры. На первом этапе выполняется оптимизацию параметров позы и камеры, на втором оцениваются параметры формы и выражения лица, на третьем происходит уточнение параметров текстуры и освещения. Производится анализ качества работы предложенного алгоритма, по результатам которого делается о его превосходстве над известными методами (DECA и классический алгоритм оценки параметров модели FLAME по RGB снимку) за счет использования карты глубины, получаемой со стереокамеры.

Третья глава посвящена описанию разработанной параметрической модели головы человека на основе нейросетевой модели представления поверхности CNeRF и двумерного нейронного рендеринга. Подробно описывается и анализируется архитектуры и модели CNeRF и модели двумерного нейронного рендеринга. Приводятся особенности обучения этих моделей в рамках единого подхода. Большое внимание

при этом автор уделяет вопросам формирования обучающих выборок, которые формируются в работе на основе открытого набора данных FFHQ и параметрической модели головы FLAME. Показывается, что такая конструкция способна синтезировать изображения, соответствующие изображениям, синтезированным при помощи модели FLAME, используя те же пространства параметров, но используя при этом неявные представления для рендеринга. Приводится описание серий экспериментов, произведённых на разработанном автором программном комплексе, в результате которых сравнивается предлагаемый подход и метод создания аватара, основанной на обычной модели NeRF. Показывается превосходство предлагаемого подхода как по точности имитации изображений, так и по скорости сходимости процедур оценки параметров модели.

В четвертой главе приводится описание разработанного метода создания параметризованного аватара головы человека и формируются методические указания по его применению. В частности обосновывается, что параметры предложенной параметрической модели возможно точно настроить для конкретного человека, используя в том числе короткую монокулярную видеопоследовательность (порядка 1000 кадров) и небольшое число итераций стохастического градиентного спуска. Это упрощает процесс формирования аватара головы человека и дает дополнительные возможности по сжатию соответствующих изображений для передачи по каналам связи с ограниченной пропускной способностью. Здесь же приводится описание алгоритма аугментации, связанного с использованием интерполяции промежуточных кадров, и приводятся разные стратегии его использования. Проводится подробное сравнительное исследование разработанного подхода с известными опубликованными решениями, а именно Neural Head Avatar, IMAvatar, PointAvatar, INSTA. Делается вывод, что предложенный метод превосходит существующие по показателям качества PSNR и LPIPS на выбранном наборе данных.

В целом материалы диссертации позволяют достаточно полно оценить объем и сложность проведенного исследования. Диссертация написана квалифицированно, литературно-техническим языком, снабжена достаточным количеством правильно оформленного иллюстративного материала. Автореферат диссертации

выполнен с соблюдением установленных требований и полностью отражает ее содержание. Следует особо отметить минимальное количество орфографических и технических ошибок в представленных материалах.

3. Научная новизна полученных результатов

Представленная работа выполнена на высоком научно-техническом уровне. Научной новизной обладают следующие, полученные автором диссертационного исследования, результаты:

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. Значимость полученных результатов для науки и производства

Представленные в диссертации результаты могут быть использованы для разработки новых систем телеприсутствия; виртуальной, дополненной, смешанной реальности; систем видеоконференцсвязи; систем создания видеоконтента и прочих системах, требующих синтеза визуально реалистичных трехмерных моделей лиц пользователей.

Практическая ценность диссертации состоит в новых возможностях описания трехмерных объектов, основанных на применении комбинации априорной информации о поверхности объекта и нейросетевой модели, неявно описывающей этот объект. Применение такой технологии позволяет обеспечить высокую точность

оценки поверхности объекта на основе обработки соответствующих RGBD изображений. Свидетельством этого может быть приведенный в диссертационной работе результат о том, что среднее расхождение между точками реального и имитированного объекта составило 0.2 мм. Немаловажно, что подобная оценка может быть выполнена по ограниченному числу кадров (менее 1000) за приемлемое время (порядка 10 минут). В связи с этим, результаты диссертационного исследования представляют большой интерес для предприятий и организаций, использующих в своей деятельности не только системы виртуальной, дополненной и смешанной реальности, но и системы построения и оценки 3D объектов в целом.

5. Степень обоснованности научных положений, выводов и рекомендаций, сформулированных в диссертации

Положения, сформулированные в диссертации Козловой Ю.Х., основываются на известных положениях фундаментальных и прикладных дисциплин: математического моделирования, вычислительных систем, компьютерного зрения, методов и алгоритмов цифровой обработки изображений. В своей работе соискатель грамотно применяет математический аппарат для решения поставленных задач, корректно интерпретирует полученные результаты, делает правильные выводы и дает обоснованные рекомендации.

Автором работы изучены и критически оценены известные результаты и теоретические положения отечественных и зарубежных ученых по рассматриваемым вопросам диссертационного исследования. Обоснованность полученных в диссертации результатов подтверждается также хорошей согласованностью с результатами других авторов. Полученные Козловой Ю.Х. научные и практические результаты соответствуют п. 1 и п. 13 паспорта специальности 2.3.8 - «Информатика и информационные процессы» (технические науки).

6.Замечания по работе

Имеются следующие замечания по диссертации.

1. В работе имеются незначительные логические несостыковки в повествовании. Например, модель Flame в первой главе рассматривается наряду со

множеством других. В выводах по первой главе автор делает фокус на нейросетевой модели представления поверхности CNeRF и двумерного нейронного рендеринга и никак ее не упоминает. Но внезапно вся вторая глава оказывается посвящена исследованию и модификации данной модели.

2. На мой взгляд, в работе недостаточное внимание уделено пояснению используемых формул и параметров в этих формулах. В частности, в формулах (2) – (5) хотелось бы видеть более подробное описание, что такое взвешивающие коэффициенты w_1 и w_2 , а также параметры p_1 и p_2 . Это несколько затрудняет понимание материала.
3. Наряду с почти безупречной общей грамотностью в работе встречаются досадные логические «ляпы». Например, на стр. 53 сказано «точность трехмерной реконструкции для предложенного алгоритма принимает настолько низкое значение, можно сделать вывод, что качество реконструкции полигональной сетки ограничено только точностью захваченного облака точек (стр. 53)».
4. Во второй главе явно отсутствуют количественные характеристики, отражающие вычислительные затраты, которые необходимы для оценки параметров Flame модели по RGBD изображению в сравнении с известными подходами.
5. Работа сопровождается значительным количеством иллюстративного материала. Только в приложении представлено 83 изображения, в каждом из которых присутствует порядка 200 результатов имитации головы человека. В реальности визуально оценить качество этих имитаций очень сложно, а всякий их анализ в работе отсутствует.
6. В работе отсутствует информации о зарегистрированном программном продукте, реализующем найденные решения. Также отсутствует информация об источнике в сети Интернет, где можно получить доступ к компилированной версии программного обеспечения автора.

Указанные замечания не умаляют достоинств представленной работы и не влияют на общую положительную оценку диссертации. Надеюсь, что часть этих замечаний будет снята во время дискуссии, а некоторые из них войдут в планы дальнейшей работы.

7. Заключение

Считаю, что в представленной на отзыв диссертационной работе решена актуальная научная задача по повышению качества и скорости формирования синтезируемых изображений-проекций аватара головы человека. Диссертация удовлетворяет требованиям «Положения о присуждении ученых степеней», соответствует паспорту специальности 2.3.8 - «Информатика и информационные процессы», а ее автор, Козлова Юлия Ханифовна, заслуживает присуждения ему ученой степени кандидата технических наук по заявленной специальности.

Заведующий кафедрой «Радиотехника,
телекоммуникации и защита информации»

ФГБОУ ВО «УлГТУ», д.т.н., доцент

Подпись д.т.н., доцента

Дементьева Виталия Евгеньевича заверяю



В.Е. Дементьев

20.11.2024

Виталий Евгеньевич Дементьев
А.Н. Ковалев

Контактная информация оппонента

Дементьев Виталий Евгеньевич, д.т.н., доцент, федеральное государственное бюджетное образовательное учреждение высшего образования «Ульяновский государственный технический университет» (ФГБОУ ВО «УлГТУ»),

заведующий кафедрой «Радиотехника,
телекоммуникации и защита информации»;

адрес организации: 432027, г. Ульяновск, ул. Северный Венец 32;

сайт организации: <http://www.ulstu.ru/>;

тел. кафедры: (8422) 77-81-23;

эл. почта: 986514@mail.ru.