

ИССЛЕДОВАНИЕ ГОЛОСОВЫХ ПОМОЩНИКОВ И ПРИМЕНЕНИЕ МЕТОДОВ TTS В ВИРТУАЛЬНОЙ РЕАЛЬНОСТИ

Вавилов М.А. (НИ ТПУ)

Научный руководитель – Видман В.В.
(НИ ТПУ)

Введение. Применение тренажеров в виртуальной реальности VR (Virtual Reality) получает все большее распространение в различных сферах, включая медицину, авиацию и промышленное обучение [1, 2]. Однако взаимодействие пользователей с VR-продуктами может вызывать определенные трудности, особенно при обучении и освоении функционала. Встраивание голосового помощника в VR-продукты может значительно упростить процесс адаптации и повысить удобство взаимодействия. Цель данного исследования заключается в изучении современных технологий TTS (Text-to-Speech), их возможностей и ограничений в контексте виртуальной реальности. Анализ существующих решений позволит определить, какие модели синтеза речи наиболее подходят для VR-продуктов.

Основная часть. Исследование посвящено изучению различных видов голосовых помощников, их архитектуры и областей применения в повседневной жизни, промышленности и в виртуальной реальности [3]. Анализ существующих решений позволит выявить сильные и слабые стороны текущих технологий, а также определить ключевые критерии для построения эффективной модели TTS.

Главной задачей исследования является изучение способов создания модели TTS, адаптированной для VR-устройств [4]. Такая модель должна не только обеспечивать высокое качество синтезированной речи, но и быть интерактивной, способной поддерживать динамичное общение с пользователем [5].

Особое внимание уделяется критериям выбора моделей синтеза речи, таким как скорость обработки, уровень натуралистичности звучания и поддержка русского языка. Также рассматриваются методы уменьшения задержек при генерации речи TTS-моделей для VR. В статье представлены сравнительные характеристики популярных моделей TTS, их архитектурные особенности, а также применимость в VR-продуктах.

Выводы. Исследование направлено на поиск эффективных решений для улучшения взаимодействия пользователей с VR-системами посредством голосовых помощников. В ходе анализа были выявлены технологии синтеза речи, которые можно применять в VR, чтобы сделать голосовое общение более естественным и комфортным. Интеграция методов TTS в VR-продуктах способствует созданию более реалистичного и удобного пользовательского опыта, делая взаимодействие с виртуальным миром более удобным и интуитивно понятным.

Список использованных источников:

1. Lizzie Waymouth. How VR and AR simulation could define the future of training offshore workers. – 2021 – URL: <https://www.nsenergybusiness.com/features/offshore-training-simulation-vr-ar/?cf-view> (дата обращения: 15.02.2025).
2. Darda P., Chitnis R. A Review on Voice Assistant Adoption in Service Sector. – 2020 – URL: https://www.researchgate.net/publication/351579569_A_Review_on_Voice_Assistant_Adoption_in_Service_Sector (дата обращения: 15.02.2025).
3. Yuqi Huang. Research on the Development of Voice Assistants in the Era of Artificial Intelligence. – 2023 – URL: <https://www.researchgate.net/publication/367077879> (дата обращения: 15.02.2025).
4. Jain K., Murphy E., Gupta D. et al. Compact Neural TTS Voices for Accessibility. – 2025

– URL: <https://arxiv.org/abs/2501.17332> (дата обращения: 15.02.2025).

5. Völkel S., Buschek D., Eiband M. et al. Eliciting and Analysing Users' Envisioned Dialogues with Perfect Voice Assistants. – 2021 – URL: <https://arxiv.org/abs/2102.13508> (дата обращения: 15.02.2025).