

УДК 004.94

3D-моделирование робота-манипулятора и система многоязычного голосового управления им

Хэ Чанфэй¹

15847735177@163.com

Лю Имин¹

liucolin20@gmail.com

Карташев Всеволод Владимирович^{2(*)}

vsevolodkartashev@gmail.com

¹ Факультет космических исследований, МГУ имени М.В. Ломоносова, Москва, Россия² ИПМ им. М.В. Келдыша РАН, Москва, Россия

Аннотация. Рассмотрено моделирование управления роботом-манипулятором с использованием фреймворков *three.js* и распознавания речи. Проведены эксперименты с голосовым управлением на нескольких языках, результаты которых представлены вместе с рекомендациями для промышленного применения. Описан формальный язык управления, куда транслируются команды на основе распознанной речи. Результаты экспериментов показывают эффективность предложенных методов для промышленных условий.

Ключевые слова: робот-манипулятор, глубокое обучение, *Single Shot Detector*, *Mask R-CNN*, сегментация изображений

Введение. Визуализация движений роботов становится всё более важной для проектирования сложных автоматизированных систем и улучшения человеко-машинного взаимодействия. В робототехнике планирование, проектирование, тестирование и отладка движений особенно сильно упрощаются при использовании трехмерной визуализации. Кроме того, с развитием технологий искусственного интеллекта способы распознавания речевых команд стали полезным средством взаимодействия человека и компьютера. В данной работе исследуется проблема трехмерного моделирования управления роботом-манипулятором с определенной кинематической схемой и методы голосового управления им.

Материалы и методы; результаты. Для спецификации модели робота-манипулятора и визуального отображения был использован фреймворк *three.js* и *VITE*. Его использование позволило осуществить моделирование движения манипулятора с несколькими степенями свободы. Для голосового управления были исследованы фреймворки для распознавания речи и фонем *Vosk*, *Google* и *allosaurus*. В работе описан формальный язык управления роботом-манипулятором. Распознанная речь на нескольких языках транслируется в команды управления роботом на формальном языке. После эти команды отправляются на вход системе визуализации через механизм *WebSocket*. Кроме того, были проведены эксперименты по распознаванию речи с использованием размеченных наборов голосовых данных [1–3].

Заключение. Результаты работы показывают, что можно успешно осуществлять голосовое управление роботом-манипулятором, подавая ему голо-

совые команды. Наибольшую пользу такой вид управления способен принести в рамках супервизорного управления. Осуществленное моделирование показывает, что данные подходы могут быть модифицированы для интеграции в робототехнические системы, используемые в производственной сфере, медицинских учреждениях и других секторах промышленности [4, 5].

Список источников

- [1] Dirksen J. *Learn Three.js: Program 3D-animations and visualizations for the web with JavaScript and WebGL*. Packt Publishing Ltd, 2023.
- [2] Manual. Three.js — JavaScript 3D Library. URL: <https://threejs.org/docs/> (дата обращения 29.11.2024).
- [3] Devlin J. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv Preprint*, 2018. URL: <https://arxiv.org/abs/1810.04805> (дата обращения 29.11.2024).
- [4] Reimers N. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. *arXiv preprint*, 2019. URL: <https://arxiv.org/abs/1908.10084> (дата обращения 29.11.2024).
- [5] Bojanowski P. Enriching word vectors with subword information. *Transactions of the association for computational linguistics*, 2017, vol. 5, pp. 135–146. URL: https://direct.mit.edu/tacL/article/doi/10.1162/tacL_a_00051/43387/Enriching-Word-Vectors-with-Subword-Information (дата обращения 29.11.2024).