

УДК 004.5:007.52:621.865.8

ОБРАЗОВАТЕЛЬНЫЙ МАНИПУЛЯТОР НА БАЗЕ НЕЙРОСЕТЕВОГО РАСПОЗНАВАНИЯ ЖЕСТОВ НА БАЗЕ КАСТОМНОЙ МОДЕЛИ YOLO11

Сафин Г.И. (ИТМО), Гриценко К.В. (ИТМО)

**Научный руководитель – инженер Гриценко К.В.
(ИТМО)**

Введение. Современная промышленная робототехника достигла высокого уровня развития, однако массовое внедрение манипуляторов сдерживается сложностью их программирования и переналадки. Традиционные методы – обучение через teach-панели или текстовые языки – требуют высокой квалификации оператора и значительных временных затрат, что особенно критично для малого бизнеса и образовательных учреждений. Существующие альтернативные подходы, такие как тактильные устройства (джойстики, 3D-мышь) или специализированные системы жестового ввода (Leap Motion, Kinect), либо не решают проблему точного позиционирования, либо имеют высокую стоимость. Таким образом, актуальной является задача создания интуитивного и доступного интерфейса управления, сочетающего простоту жестового управления с точностью аналитических расчётов кинематики.

Основная часть. В работе предлагается архитектура системы, объединяющая три ключевых компонента в едином контуре реального времени. Первый компонент – модуль нейросетевого распознавания жестов на базе кастомной модели YOLO11. Модель обучена на авторском датасете изображений кисти человека в различных рабочих позах и условиях освещения. Она детектирует 21 ключевую точку кисти, что позволяет определять расстояние между большим пальцем и мизинцем, интерпретируемое как степень раскрытия хвата. Использование стандартной веб-камеры (разрешение 1280×720, частота 30 FPS) вместо специализированных сенсоров глубины обеспечивает низкую стоимость и доступность решения.

Второй компонент – вычислительное ядро обратной кинематики, реализующее аналитическое решение для SCARA-подобной кинематической схемы манипулятора с тремя степенями свободы (два вращательных звена и хват). В отличие от итеративных численных методов, аналитическое решение гарантирует детерминированность, высокую скорость вычислений и отсутствие проблем сходимости. Алгоритм преобразует координаты цели, задаваемые положением кисти оператора, в углы поворота сервоприводов с задержкой управления не более 150 мс, что позволяет работать в реальном времени.

Третий компонент – клиент-серверное программное обеспечение на Python. Серверная часть, развёрнутая на одноплатном компьютере Raspberry Pi 5, отвечает за захват видеопотока, распознавание жестов и расчёт кинематики. Клиентское приложение на PyQt6 предоставляет графический интерфейс для мониторинга и управления: отображение видеопотока с наложением ключевых точек, схему манипулятора, графики изменения углов сочленений, тепловую карту рабочей зоны, систему логирования и модуль генерации отчётов в

формате PDF. Открытая модульная архитектура позволяет адаптировать систему под конкретные задачи.

Механическая часть прототипа спроектирована в САПР КОМПАС-3D и изготовлена методом FDM-печати из ABS-пластика, что обеспечивает прочность, лёгкость и ремонтпригодность. Управление приводами (сервоприводы MG996R, шаговый двигатель NEMA17) осуществляется через микроконтроллер Arduino Uno, получающий команды от Raspberry Pi по последовательному интерфейсу.

Выводы. Разработанный прототип успешно прошёл испытания: время отклика не превышает 150 мс, частота обработки видеопотока – до 30 кадров/с. Практическая значимость проекта заключается в снижении порога вхождения в робототехнику: время обучения оператора сокращается с нескольких недель до 30 минут, а себестоимость системы при мелкосерийном производстве оценивается в 28–35 тыс. рублей, что значительно ниже промышленных аналогов. Результаты могут быть внедрены в образование (учебно-лабораторные комплексы), малый бизнес (быстрое прототипирование, автоматизация) и научные исследования (отработка алгоритмов человеко-машинного взаимодействия). Перспективы развития включают интеграцию с дополненной реальностью, расширение жестового словаря, адаптивные алгоритмы и облачный сервис для сбора данных.

Список использованных источников:

1. Arduino: официальный сайт [Электронный ресурс]. – URL: <https://www.arduino.cc/> (дата обращения: 02.02.2025).
2. PyTorch Documentation [Электронный ресурс]. – URL: <https://pytorch.org/docs/stable/index.html> (дата обращения: 15.01.2025).
3. OpenCV: Open Source Computer Vision Library [Электронный ресурс]. – URL: <https://opencv.org/> (дата обращения: 10.01.2025).
4. YOLO: Real-Time Object Detection [Электронный ресурс]. – URL: <https://pjreddie.com/darknet/yolo/> (дата обращения: 05.02.2025).
5. PyQt6 Documentation [Электронный ресурс]. – URL: <https://www.riverbankcomputing.com/static/Docs/PyQt6/> (дата обращения: 20.01.2025).
6. Кинематика и динамика роботов-манипуляторов: аналитические методы. / Ушаков Д.М. – Новосибирск : Наука, 2020. – 264 с.
7. Глубокое обучение и компьютерное зрение на Python. / Пьенг А. – Москва : Диалектика, 2022. – 480 с.
8. Компьютерное зрение на Python. / Кирьянов Д.В. – Санкт-Петербург : БХВ-Петербург, 2023. – 416 с.