

Введение. В современной системе развития робототехники компьютерное зрение и технологии управления роботизированными манипуляторами играют важную роль. Обеспечение точности и эффективности захвата манипулятора по-прежнему остается актуальной темой исследований. Однако традиционные методы управления манипулятором обычно предполагают его перемещение между фиксированными точками, что ограничивает его адаптацию к изменяющимся условиям окружающей среды и позволяет выполнять только повторяющиеся механические операции. Компьютерное зрение, являясь «глазами» робота, значительно повышает его адаптивность. В последние годы оно достигло значительного прогресса в таких областях, как обнаружение объектов, семантическая сегментация и инстанс-сегментация. Использование глубоких нейронных сетей (таких как PointNet[1], PointPillar[2], YOLO[3], FCN[4]) для семантического распознавания окружающей среды стало классическим методом в области роботизированного обнаружения объектов. В то же время развитие алгоритмов оценки позы, таких как триангуляция и итеративный алгоритм ближайшей точки (ICP)[5], привело к значительному прогрессу в технологии захвата манипуляторов.

В данной работе исследуется алгоритм захвата манипулятора на основе технологий компьютерного зрения с целью повышения точности захвата и улучшения взаимодействия человека с роботом. Предлагается алгоритм визуального захвата, основанный на слиянии Fully Convolutional Network (FCN) и ICP. Перед выполнением захвата манипулятор использует FCN для семантической сегментации RGB-D изображений, что позволяет выполнить пиксельное позиционирование объекта. Затем с помощью ICP определяется ориентация и положение объекта, после чего манипулятор осуществляет захват и размещение. Ключевым новшеством является объединение адаптивности глубокого обучения с надежностью традиционного геометрического сопоставления, что делает предложенный метод особенно эффективным для точного захвата жестких промышленных объектов.

Основная часть. Для реализации алгоритма захвата манипулятора на основе компьютерного зрения необходимо решить следующие ключевые проблемы:

- 1) Сначала необходимо выполнить калибровку между рукой робота и камерой. Калибровка является ключевым шагом в определении точного позиционного соотношения между камерой и рукой робота. Благодаря калибровке роботизированная рука может гарантировать точность захвата. Улучшить показатели успешности захвата и эффективность производства. В то же время была выбрана система Eye-in-Hand, которая устанавливает камеру непосредственно на конце роботизированной руки, динамически регулирует угол обзора посредством синхронного движения в реальном времени и обеспечивает высокоточную работу и гибкую реакцию при выполнении таких задач, как динамический захват.[6]
- 2) Подготовка набора данных для обучения FCN. В рамках использования системы Eye-in-Hand возможно непосредственно задействовать движение роботизированной руки для сбора и аннотирования данных. FCN представляет собой модель глубокого обучения для сегментации на уровне пикселей, которая преобразует объекты на изображении в информацию о местоположении, подходящую для точного восприятия.[7]
- 3) Наконец, необходимо определить точную позу объекта, чтобы облегчить дальнейшую работу по захвату с помощью манипулятора. Поскольку на раннем этапе использовалась камера глубины, полное визуальное восприятие включает захват исходного RGB-изображения сцены и соответствующего облака точек глубины. Для подгонки идеальных данных облака точек с фактическими данными используется алгоритм ICP. Когда точность подгонки достигает максимума, получается фактическая поза объекта. На последующих этапах роботизированной руке достаточно выполнить захват, основываясь на данных о позе

и положении, полученных ранее.

Выводы. Этот алгоритм представляет собой многогранную интеграцию технологий компьютерного зрения, глубокого обучения и управления роботизированной рукой. На его основе роботизированная рука может захватывать целевой объект с высокой точностью и эффективностью, а также демонстрирует хорошую устойчивость к изменяющимся условиям. В будущем этот алгоритм может сыграть значительную роль в промышленной сфере.

Список использованных источников:

1. Qi C. R. et al. Pointnet: Deep learning on point sets for 3d classification and segmentation //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2017. – С. 652-660.
2. Lang A. H. et al. Pointpillars: Fast encoders for object detection from point clouds //Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. – 2019. – С. 12697-12705.
3. Redmon J. et al. You only look once: Unified, real-time object detection //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2016. – С. 779-788.
4. Long J., Shelhamer E., Darrell T. Fully convolutional networks for semantic segmentation //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2015. – С. 3431-3440.
5. Besl P. J., McKay N. D. Method for registration of 3-D shapes //Sensor fusion IV: control paradigms and data structures. – Spie, 1992. – Т. 1611. – С. 586-606.
6. Horaud R, Dornaika F. Hand-eye calibration[J]. The international journal of robotics research, 1995, 14(3): 195-210.
7. Levine S, Pastor P, Krizhevsky A, et al. Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection[J]. The International journal of robotics research, 2018, 37(4-5): 421-436.