

Навигация наземного робота с использованием глубокого обучения с подкреплением в иммерсивной среде

Л.А. Жердева¹, Д.А. Жердев^{1,2}, Е.Ю. Минаев¹

¹Самарский национальный исследовательский университет им. академика С.П. Королева, Московское шоссе 34а, Самара, Россия, 443086

²Институт систем обработки изображений РАН - филиал ФНИЦ «Кристаллография и фотоника» РАН, Молодогвардейская 151, Самара, Россия, 443001

Аннотация

В статье представлены исследования метода глубокого обучения с подкреплением для задачи автономного позиционирования малого робота в имитационной среде. В проведенных экспериментах используется открытый игровой движок Unreal Engine с моделированием физически адекватной 3D модели сцены с препятствиями. Получаемые изображения виртуальной моделью робота в среде симуляции подаются в нейронную сеть для определения необходимой в работе информации о направлении цели и препятствиях в окружающей среде и последующей возможности обучения реального робота с так называемым подкреплением. В рамках данной работы нами исследуется способность агента в процессе обучения свободного движения без взаимодействия с другими статичными или движущимися объектами на сцене.

Ключевые слова

Обучение с подкреплением, глубокое обучение, DQN, робот, моделирование

1. Введение

Повсеместное использование автономных систем и роботов значительно выросло за счёт разнообразия сценариев использования, в том числе для задачи доставки и эксплуатации на опасных объектах [1]. Обычно комплект навигации состоит из множества модулей, в том числе одновременной локализации и сопоставления, используя для этого наблюдения с сенсорных датчиков и камер. Для успешного применения полностью автономных роботов в последние годы всё чаще используется передача данных обучаемого агента из моделируемой среды к реальному физическому устройству, где глубокое обучение используется для прогнозирования движения роботов [2]. Основной проблемой навигации робота является незнание окружающей среды, в частности, если высока вероятность взаимодействия агента с движущимися объектами. Основанное на человеческом способе обучения, обучение с подкреплением успешно может решить данную проблему. Подобные модели учатся в окружающей среде на основании оценки успешного и неуспешного автоматизированного пилотирования, где на каждом шаге обучения агент наблюдает за окружающей средой и получает текущее состояние и далее принимает действие, за которое получает некоторое вознаграждение из функции вознаграждения. Наградой выступает скалярная величина, информирующая о пользе или неудобстве такого действия, а целью агента является максимизация дисконтированного вознаграждения [3]. Как частный и известный метод обучения с подкреплением, в случае отсутствия знаний об окружающей среде – это Q-обучение. В работе [4] показано, что возможно успешное использование комбинации Q-обучения со сверточной нейронной сетью (CNN) – глубокое Q-обучение (DQN), где изображение является входным параметром CNN, а функция, определяющая действие - выходным параметром. Использование DQN показало внушительные результаты, превосходящие человека, для виртуальных агентов [5]. В данной работе представлены результаты исследования симуляции самообучающейся навигации и избегания столкновений со статическими и подвижными препятствиями для наземного робота в сложной окружающей обстановке.

2. Обучение агента в синтетической среде

В качестве основы среда трехмерной симуляции для экспериментов была реализована с использованием открытого для разработок трехмерных сцен и сценариев игрового движка Unreal Engine. Сцена для обучения агента представляет собой реалистичную комнату фиксированной длины, в которой произвольно размещались различные статические препятствия, такие как перегородки, коробки, стулья и тд. В качестве агента выступает виртуальный наземный робот, на котором располагается виртуальная фронтальная камера, фиксирующая двухмерное изображение на виртуальную матрицу. Выходной сигнал DQN представляет собой дискретный четырехмерный вектор: двигаться вперед или назад, поворот вправо или влево. Для отображения состояния на соответствующие Q-значения использована свёрточная нейронная сеть, а в качестве основы для построения программы обучения с подкреплением используется фреймворк TensorFlow для глубокого обучения. Каждый эпизод обучения агент начинает двигаться произвольно из центра координат до некоторой точки назначения, получая изображение и выбирая действие. Были проведены два типа экспериментов: обучение в комнате только со статическими объектами, и в комнате с 1-5 подвижными объектами. Эпизод обучения заканчивался, когда агент достигал пункта назначения, за что получал награду, в противном случае при столкновении с препятствиями агент получал отрицательную награду.

3. Заключение

Продемонстрированы результаты модельных исследований метода глубокого обучения с подкреплением для задачи избегания столкновений с неподвижными и динамичными объектами в трёхмерной среде. Представлен анализ поведения и политики избегания движущихся объектов агентом, обученном в среде только со статичными объектами, и агентом, обученном в среде с наличием динамичных объектов. Рассмотрена возможность доработки и развёртывания обученной системы навигации на реальном прототипе мобильного наземного робота.

4. Литература

- [1] Paelke, V. Augmented reality in the smart factory: Supporting workers in an industry 4.0 environment / V. Paelke // Proceedings of the IEEE emerging technology and factory automation (ETFa). – 2014. – P. 1-4.
- [2] Wei, X.L. An Improved Method Based on Deep Reinforcement Learning for Target Searching / X.L. Wei, X.L. Huang, T. Lu, G.G. Song // Proceedings of the 4th International Conference on Robotics and Automation Engineering (ICRAE). – 2019. – P. 1-5.
- [3] Kober, J. Reinforcement learning in robotics: A survey / J. Kober, J.A. Bagnell, J. Peters // The International Journal of Robotics Research. – Vol. 32(11). – 2013. – P. 1238-1274.
- [4] Mnih, V. Human-level control through deep reinforcement learning / V. Mnih, K. Kavukcuoglu, D. Silver, A.A. Rusu, J. Veness // Nature. – 2015. – Vol. 518. – P. 529-533.
- [5] Lample, G. Playing FPS Games with Deep Reinforcement Learning / G. Lample, D.S. Chaplot // Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17). – 2017. – P. 2140-2146.