

Қазақстан Республикасы Тәуелсіздігінің 30 жылдығына арналған «Сейфуллин оқулары – 17: «Қазіргі аграрлық ғылым: цифрлық трансформация» атты халықаралық ғылыми – тәжірибелік конференцияға материалдар = Материалы международной научно – теоретической конференции «Сейфуллинские чтения – 17: «Современная аграрная наука: цифровая трансформация», посвященной 30 – летию Независимости Республики Казахстан.- 2021.- Т.1, Ч.4 - С.212-215

## **ПОДХОД К АЛГОРИТМАМ ОБУЧЕНИЯ НАВИГАЦИОННОЙ СИСТЕМЫ ДЛЯ МОБИЛЬНЫХ РОБОТОВ**

*Вьюшков.Р.*

В данной работе представлена самосовершенствующаяся система непрерывного обучения для мобильного робота, ориентирующегося в различных средах. Классические статические методы навигации требуют специфической настройки системы, например, от экспертов-людей, или же, могут повторять свои ошибки независимо от того, сколько раз они использовались в одной и той же среде. В данной статье рассматривается «Пожизненное навигационное обучение», которое улучшает навигационное поведение мобильного робота исключительно на основе его собственного опыта и сохраняет способность робота ориентироваться в предыдущих средах после обучения в новых. ПНО реализуется и тестируется полностью на борту физического робота с ограниченным объемом памяти и вычислений.

Классические мобильные роботы предназначены для адаптации к различным навигационным средам путем настройки базовой навигационной системы на месте, например при помощи калибровки датчиков или настройки параметров. Однако без применения и корректировки экспертных знаний ненастроенная система может повторять одни и те же ошибки (например, застрять в одном и том же узком месте) даже несмотря на то, что устройство несколько раз и более использовалось в одной и той же среде.

Недавний успех в использовании машинного обучения для навигации мобильных роботов указывает на потенциал улучшения навигационных характеристик на основе прошлого опыта робота в той же среде [1]. Однако при столкновении с различными навигационными средами методы обучения не могут хорошо распространяться на новые сценарии: роботы должны заново учиться ориентироваться в новых средах. Что еще более важно, изученная система склонна к катастрофическому забыванию, что заставляет робота забывать то, что было изучено в предыдущих средах [2]. Робот способен определять свои неоптимальные действия и учиться на них. Сталкиваясь с различными навигационными средами, робот способен учиться адаптироваться к новым средам, не забывая при этом, как ориентироваться в предыдущих окружающих его средах. Основными рассматриваемыми методами этой статьи являются:

- Стратегия самосовершенствования, которая позволяет динамически повышать производительность навигации;
- Схема обучения на протяжении всей жизни, которая позволяет роботу ориентироваться в новых условиях, не забывая при этом о предыдущих;
- Внедрение системы непрерывного обучения навигации полностью на борту физической роботизированной платформы с ограниченным объемом памяти и вычислениями.

Классические навигационные системы предназначены для применения в самых разнообразных средах, но обычно они статичны, работают при фиксированном наборе заранее заданных параметров и поэтому таким системам не хватает способности

совершенствоваться с опытом и адаптироваться к конкретной среде. Настройка параметров - это современная практика, которая требует от человека-эксперта интуиции, опыта и проб и ошибок. В большинстве случаев настройка требует человеческих знаний, и также дополнительной калибровки датчиков или демонстрации работы навигации. Кроме того, после настройки статическая навигационная система теряет способность к дальнейшему совершенствованию с большим опытом или адаптации к новым условиям.

В отличие от этого, предлагаемое Пожизненное обучение навигации не требует человеческих знаний и может динамично улучшаться с увеличением навигационного опыта при столкновении с новыми средами.

Навигация на основе обучения: Методы машинного обучения, основанные на данных получаемые в реальном времени, также широко применяются для решения навигационных задач. Что касается физической навигации роботов, подходы к обучению обычно настраиваются методом проб и ошибок, используя обучение с подкреплением. В то время как эти методы обучения позволяют улучшить конкретную среду с повышенным опытом навигации, если робот должен быть помещен в несколько сред последовательным образом, что часто встречается в реальной навигации, методы обучения могут плохо обобщаться и могут легко забыть прошлые знания. Метод НПО рассматривает, как предотвратить забывание и увеличить обобщение.

Пожизненное/непрерывное контролируемое обучение: пожизненное или непрерывное обучение изучает проблему обучения на постоянной основе. В последнее время был достигнут значительный прогресс в непрерывном обучении с помощью нейронных сетей.

В тестах участвовало наземное устройство, используемое как для имитационных, так и для физических экспериментов. Робот оснащен лазерным сканером для восприятия окружающих препятствий и работает под управлением базовой операционной системы робота (ROS) `move_base navigation stack`.

Работа ПНО сначала изучается с помощью обширных имитационных испытаний. На борту устройства имеется лазерный сканер LMS111, обеспечивающий 270° 720-мерное лазерное сканирование. Три смоделированные навигационные среды показаны на рис. 1, где робот перемещается от фиксированного старта к фиксированной цели в окружающей среде. В моделируемой навигационной среде №1, требуется быстрая реакция для обхода препятствий; В моделируемой навигационной среде № 2 робот должен поддерживать медленный и последовательный темп, чтобы проехать через узкий проход; В моделируемой навигационной среде №3 требуется, чтобы робот замедлился, для того чтобы сделать резкий поворот и плавно войти в другую комнату.

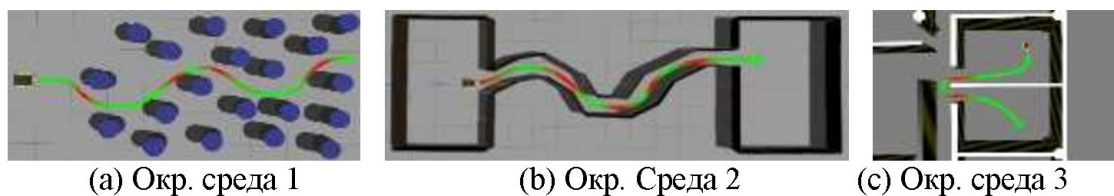


Рисунок 1 - Моделируемые навигационные среды

Для проведения эксперимента была использована нейронная сеть из трех скрытых слоев с 64 скрытыми нейронами в каждом, чтобы вычислить линейную и угловую скорость на основе входных данных лидара и локальной цели. После обучения происходит оценка навигационных характеристик с точки зрения времени прохождения в текущей и предыдущей навигационных средах. На рисунке 2 отображены результаты:

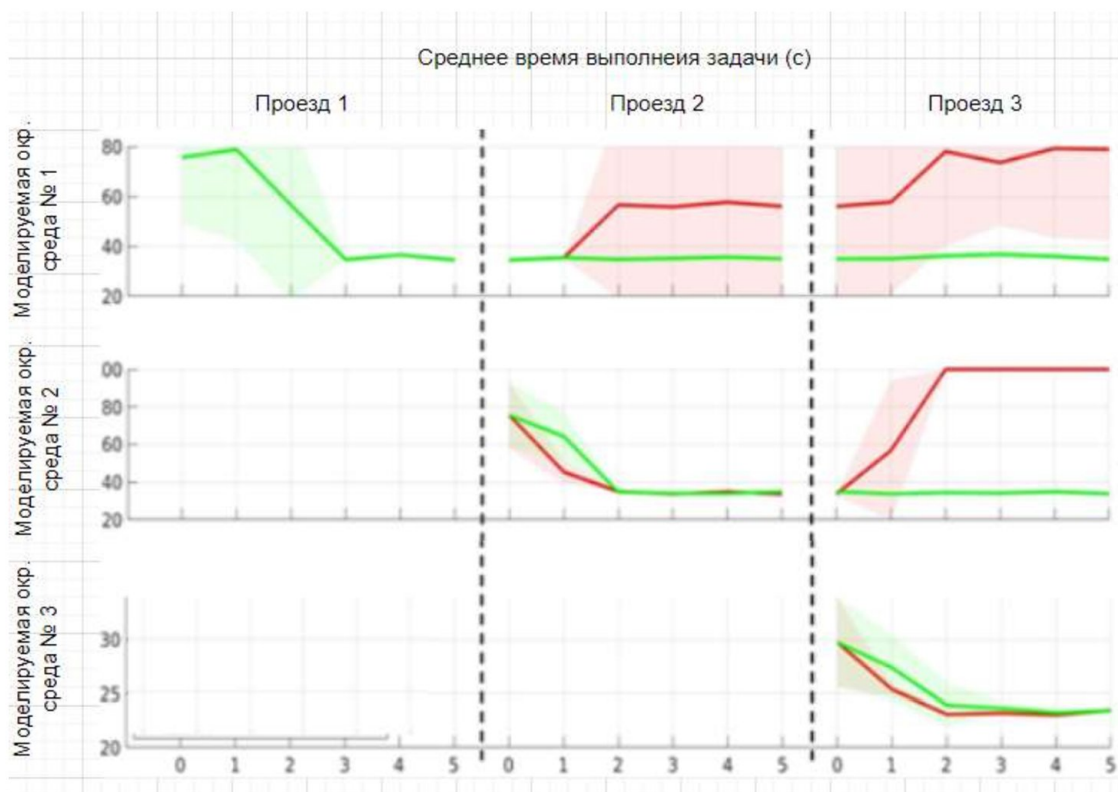


Рисунок 2 - Результаты эксперимента: навигационные характеристики в отношении увеличения опыта

Политика выполняется трижды, в результате чего в общей сложности проводится 198 оценочных испытаний. Мы сообщаем среднее и стандартное отклонение производительности на рисунке 2. Если робот не в состоянии достичь цели, например застрянет, дается штрафное время в 100 секунд. В каждой навигационной среде ПНО способен сократить время прохождения с увеличением опыта обучения. В разных средах ПНО изучает новые среды с увеличением объема данных, избегая при этом катастрофического забывания предыдущих. Последовательное обучение может улучшить навигационные характеристики данной среды при представлении данных из этой конкретной среды быстрее, чем ПНО. Однако производительность навигации в предыдущих средах ухудшается с увеличением опыта работы в текущей среде. Катастрофическое забывание проявляется в отклонении красной линии (Последовательное обучение) от зеленой линии.

Опираясь на исходную статическую модель, основанную на выборке, прогнозирующую политику управления, которая улучшается с увеличением навигационного опыта, робот способен самоидентифицировать неоптимальные действия, искать аналогичные сценарии, в которых выполняются оптимальные действия, учиться на этих данных и постоянно улучшать навигацию. Кроме того, в условиях мультисреды ПНО способно адаптироваться к новым условиям, не забывая при этом о предыдущих окружающих средах. Обширные имитационные испытания проводятся для проверки способности ПНО к обучению в окружающей среде и кросс-среде. Весь алгоритм также реализован и протестирован на ограниченных вычислительных ресурсах на борту физического робота и работает в режиме реального времени, не требуя каких-либо внеплатных вычислений.

#### Список литературы

1. G. Kahn, A. Villaflor, B. Ding, P. Abbeel, and S. Levine, "Self-supervised deep reinforcement learning with generalized computation graphs for robot navigation," in 2018 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2018, pp. 1–8.

2. R. M. French, “Catastrophic forgetting in connectionist networks,” Trends in cognitive sciences, vol. 3, no. 4, pp. 128–135, 1999

3. Madridano, Á., Al-Kaff, A., Martín, D., de la Escalera, A. Trajectory planning for multi-robot systems: Methods and applications (2021) Expert Systems with Applications, 173, статья № 114660, . 1) <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85100882939&doi=10.1016%2fj.eswa.2021.114660&partnerID=40&md5=fc2dca1cc6dcc7d3ccaea46d816a17c7> DOI: 10.1016/j.eswa.2021.114660

*Научный руководитель Исмаилова А.А., PhD*