

Мобильный робот в нестационарной среде

Т.И. Чачхиани, М.Г. Серова

Нижегородский государственный университет им. Н.И. Лобачевского

Аннотация: Разработан исследовательский программный комплекс для формирования поведения мобильного робота в нестационарной среде на базе робота e-puck. Комплекс включает в себя методику дообучения робота, кластерный анализ и модифицированный алгоритм Q-learning, обеспечивающий движение робота к цели.

Ключевые слова: автономный мобильный робот e-puck, алгоритм Q-Learning, кластерный анализ, обучение, распознавание, цифровые методы обработки видеоизображений.

Введение

Математические методы применяются в самых различных прикладных областях [1 – 4]. Использование методов теории распознавания образов для идентификации и достижения объектов мобильным роботом является перспективным направлением. В основе теории распознавания образов лежат такие фундаментальные понятия как «обучение» и «распознавание». Это позволяет использовать эти методы в анализе видео-информации с камеры робота и в обучении робота распознаванию объектов.

Для обеспечения интеллектуальности робота, необходимо наделить его возможностью приспосабливаться к окружающей среде. С этой целью используются различные алгоритмы обучения. В тех случаях, когда у робота отсутствует или сильно ограничена информация о структуре окружающей среды, особый интерес могут представлять алгоритмы обучения с подкреплением.

Исследованию алгоритмов обучения и обучения с подкреплением посвящено достаточно публикаций [5 – 11]. Вместе с тем, не смотря на большое количество предлагаемых алгоритмов, остается множество нерешенных проблем:

- В целом весьма трудно определить все состояния среды и возможные действия системы в реальном мире.

- Проблема частичной наблюдаемости. В реальных задачах робот часто не знает, в каком именно состоянии он окажется после применения выбранного действия.
- Проблема структуры вознаграждения. Задача состоит в том, как лучше распределить вознаграждение между различными вариантами отклика. При этом ответ среды может быть многомерным.
- Проблема в оценке адекватности соотношения базы знаний робота и окружающего мира - в какой момент он может перестать исследовать среду и просто использовать имеющиеся знания.
- Нестационарность среды. Это фундаментальная проблема, которая до сих пор слабо исследована.

В данной работе рассматривается обучение мобильного робота в нестационарной среде. В качестве интеллектуального робота выступает автономный мобильный робот e-ruck (рис.1).

E-ruck

- Диаметр: 70 мм
- Высота: 50 мм
- Вес: 200 г
- Максимальная скорость: 13 см/с
- Время автономной работы: 2 часа непрерывного движения
- Процессор: dsPIC 30 CPU @ 30 MHz (15 MIPS)
- **Память: 8 KB RAM; 144 KB Flash**
- 2 шаговых электродвигателя, вращающих колёса
- 8 инфракрасных излучателей и бесконтактных датчиков (TCRT1000)
- **Цветная камера 52x39 (реальное разрешение 640x480)**
- 8 светодиодов на внешнем кольце + 1 светодиод для подсветки корпуса + 1 фронтальный светодиод
- 3D акселерометр
- 3 микрофона
- 1 динамик



Рис. 1. Автономный мобильный робот e-ruck.

Управление роботом предполагается осуществлять посредством постоянной передачи информации с сенсоров на вычислительный комплекс

(ПК) и обратной передачи управляющих сигналов, как результата анализа всего предшествующего поведения робота. Создается поведенческая модель обучающегося робота. Эта модель включает в себя многоэтапный анализ среды, отслеживание изменений в ней, выбор способа поведения в текущей ситуации.

Постановка задачи

Цель работы – создание программного комплекса для исследования проблемы обучения на базе мобильного робота e-ruck.

Была сформулирована задача обучить робота анализировать окружающую среду с целью нахождения в ней определенных объектов. Окружающая среда не является стационарной - объекты могут меняться со временем как внешне, так и менять свою позицию. Поэтому возникает необходимость разработки алгоритма постоянного дообучения для распознавания новых объектов.

В качестве основного источника информации для робота являются данные с камеры. Доступное с камеры робота изображение имеет разрешение 52x39, глубину цвета 8 бит. Все значимые объекты сцены помечаются специальными метками. В качестве таковых используются различные знаки заранее определенного цвета (например, синего).

Целевые объекты, которые требуется найти, задаются пользователем. В процессе работы цель может меняться, при этом возможно переключение между известными целями, добавление новых. Если при этом объект уже встречался, алгоритм должен учитывать предыдущий опыт.

Идея и реализация задачи

Робот e-ruck осуществляет деятельность в нестационарной среде с целью получения максимальной выгоды. Вознаграждение он получает,

достигая поставленной цели. Поэтому почти весь его жизненный цикл можно представить в следующем виде:

1. предварительная обработка информации с сенсоров;
2. учет вознаграждения среды, выработка дальнейших действий;
3. действие;
4. переход к пункту 1.

Обработка информации с сенсоров

На каждой итерации снимается изображение с камеры робота и данные с ИК-датчиков окружения. Основное внимание на данном этапе уделяется обработке изображения с камеры. На рис.2 представлены последовательные шаги анализа изображения.



Рис. 2. Последовательные шаги анализа изображения.

Для реализации шагов 1 – 4 использовалась библиотека OpenCV с целью увеличения производительности вычислений и облегчения процедуры создания комплекса.

Распознавание изображения с использованием кластерного анализа

Так как обучающая выборка отсутствует и количество будущих объектов в ходе эксперимента заранее неизвестно, использовался алгоритм кластерного анализа. Он умеет принимать на вход изображение (метку

объекта), оценивать принадлежность его одному из известных кластеров и модифицировать свою базу данных соответствующим образом. В результате его работы определяется номер кластера из базы данных алгоритма для метки объекта.

Алгоритм распознавания построен на определении сходства между изображениями и формировании в процессе обучения кластеров меток для каждого объекта. Таким образом, из похожих меток постепенно формируются эталонные изображения для соответствующих объектам кластеров.

Если метка сходна с имеющимся эталоном, то считается, что объект встречался ранее и соответствующее действие робота уже было определено. Иначе создается новый кластер. Таким образом, происходит дообучение. Количество новых объектов практически неограниченно. Новая метка становится эталонным изображением этого кластера.

Для того чтобы определить принадлежность новой метки какому-либо из уже имеющихся кластеров, определяется «расстояние» между эталонной меткой и исследуемой меткой. При определении «расстояния» учитывается, что расстояние между пикселями равно модулю разности их яркостей.

Если «расстояние» меньше пороговой величины, то метка считается принадлежащей найденному кластеру. Кроме того, проверяется не принадлежит ли метка еще какому-либо кластеру. Если такой найдется, то его следует объединить с первым найденным с целью предотвращения дублирования. Происходит коррекция эталонной метки кластера на основе новой метки, отнесенной к этому кластеру.

В качестве примера можно рассмотреть тестовое изображение и его монохромный аналог (после пороговой обработки), готовый к анализу (рис.3).

Можно увидеть, что текст написан не идеально. Это отражает возможные искажения на реальном изображении, получаемом с камеры. Обход монохромного аналога изображения (рис. 3) происходит построчно слева направо, сверху вниз.

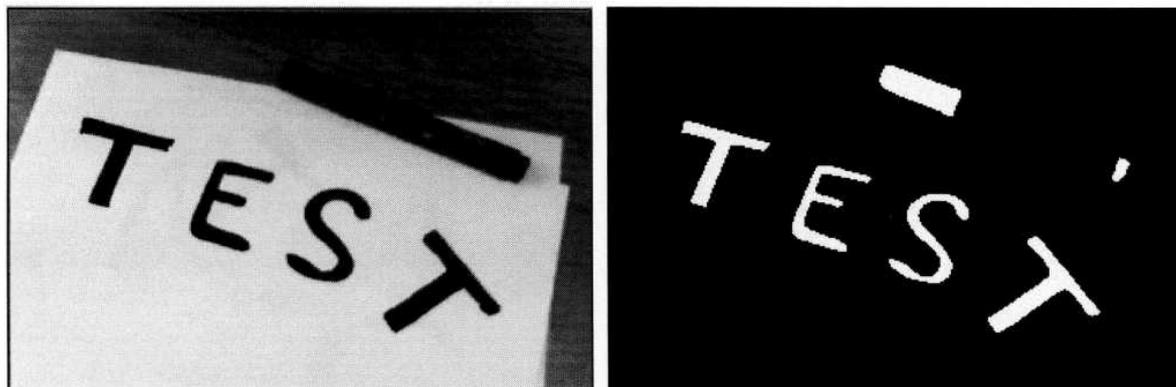


Рис. 3. Исходное и монохромное изображения.

На рис. 4 показан процесс образования кластеров. Сначала выделяется верхний элемент (колпачок маркера), классифицируются как объект кластера №1 (рис. 4а) и становится эталоном этого кластера.

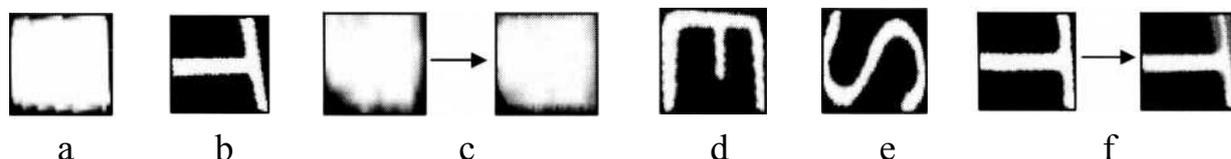


Рис. 4. Процесс образования кластеров: а – эталон кластера №1; б – эталон кластера №2; с – коррекция эталона кластера №1; д – эталон кластера №3; е – эталон кластера №4; ф – коррекция эталона кластера №2.

При выделении левой буквы Т «расстояние» до имеющегося кластера №1 составило 433707 единиц яркости, то есть порядка 68% от максимально возможного изменения. Порог разницы в 20% не позволяет их считать одним кластером, поэтому создается кластер №2 (рис. 4б) и буква Т становится эталоном этого кластера.

Правая точка от маркера классифицируется как объект кластера №1 в силу своего сходства с эталонным этого кластера: «расстояние» между ними

составило 93948 единиц яркости (разница в 14% допустима). Производится коррекция эталона кластера №1 (рис. 4с).

Следующей выделяется буква Е. При этом «расстояние» до эталона кластера №1 - 405419 (63%), до эталона кластера №2 - 271003 (42%). Так как «расстояния» велики, создается кластер №3 и буква Е становится эталоном этого кластера (рис. 4d).

Аналогично происходит анализ буквы S: «расстояние» до эталона кластера №1 - 388234 (60%), «расстояние» до эталона кластера №2 - 258467 (40%) и «расстояние» до эталона кластера №3 - 254856 (39%). Создается новый кластер №4 и буква S становится эталоном этого кластера (рис. 4е).

Иная ситуация при анализе левой нижней буквы Т. «Расстояние» до эталона кластера №2 - 86320 (13% от максимума). Благодаря этому буква Т распознается как объект кластера №2; вносится коррекция в эталон кластера №2 (рис. 4f).

На рис.5 представлены исходное и проанализированное изображение с указанием номеров кластеров для элементов изображения.

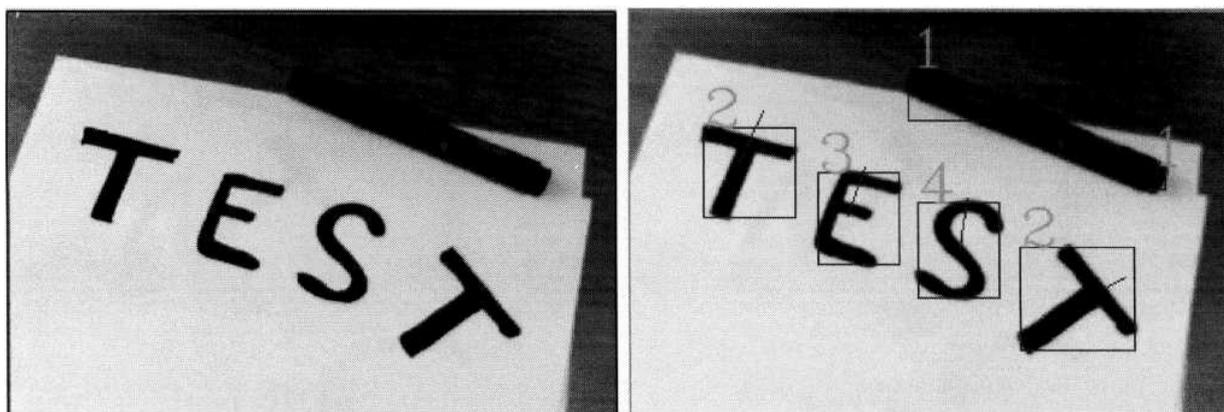


Рис. 5. Исходное и проанализированное изображения.

Соединение информации обо всех объектах на исходном изображении

Данная операция позволяет рассматривать каждый кадр с видеокамеры как отдельное состояние робота в окружающей среде: в результате

получаем некое обобщение текущих данных для принятия решений о дальнейших действиях.

В каждом новом состоянии повторяются 3 операции:

1. Робот получает данные с сенсоров и производит их предварительную обработку: сглаживает, выделяет и поворачивает метки в кадре.
2. После этого происходит распознавание и классификация меток.
3. На основе полученной информации принимается решение о том, что робот будет делать в текущем состоянии.

Принятие решений. Использование модифицированного алгоритма обучения с подкреплением.

Подход на основе методов обучения с подкреплением достаточно распространен в связи со своей логичностью и простотой реализации. Но также имеет ряд ограничений - в частности имеются серьезные требования в виде быстрой работы с базой данных и, как следствие, её размерам.

В основе метода лежит классический алгоритм обучения с подкреплением Q-learning [6]. Он обеспечивает движение робота к цели. Алгоритм работает на основе матрицы ожидаемого вознаграждения Q , изменяемой со временем в зависимости от получаемого вознаграждения.

Получаемое вознаграждение определяется самим роботом в зависимости от текущего состояния окружающей среды. Цель, которую робот должен достичь в текущий момент, указывается пользователем. Алгоритм классифицирует целевую метку и анализирует текущее состояние среды на наличие искомого объекта в ней.

Зная текущее состояние среды, робот выбирает действие, которое, как ожидается, принесет максимальное вознаграждение. Вознаграждение выдается, если целевая метка находится в определенной области кадра видеокамеры и имеет определенный размер. Изначально матрица Q задается

случайно. Кроме того, на каждом шаге значения матрицы меняются таким образом, что матрица дает оптимальное управление в любой ситуации. В случае робота e-risk, возможные действия – прямолинейное движение или поворот (налево/направо).

На основе полученного вознаграждения изменяются значения матрицы вознаграждений Q . Однако известно, что при смене цели использование старой матрицы Q невозможно - её структура позволяет эффективно искать лишь один объект, и при смене задачи алгоритму придется переучиваться. Поэтому была введена следующая модификация.

Схема модифицированного алгоритма представлена на рис.6.

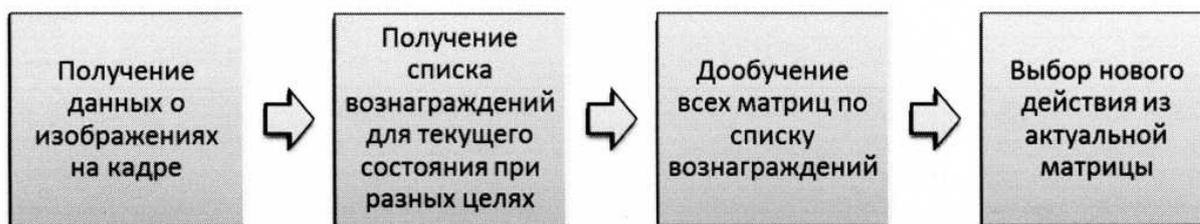


Рис. 6. Схема модифицированного алгоритма.

Для преодоления проблемы с несколькими целями, алгоритм использует несколько матриц вознаграждения Q . При возникновении новой цели алгоритм создает новую матрицу вознаграждений. При этом модификации подвергаются все матрицы Q , независимо от цели, которую ищет робот. Окончательное решение о действии робот принимает на основе матрицы, соответствующей текущей цели.

Таким образом, при переключении между матрицами, может решаться задача поиска разных объектов без потери уже обученных матриц. Следовательно, при появлении новой цели можно избавиться от необходимости постоянного переобучения, что дает возможность работы в нестационарной среде.

Вывод

Таким образом, реализован начальный этап создания программного комплекса для формирования поведения робота в нестационарной среде.

Проведенные исследования выявили достаточно большое количество проблем. Многие проблемы возникли в связи с ограничениями в конструкции робота e-risk, на основе которого строилась система. Кроме проблем нестационарной среды, существует проблема малого объема памяти робота, ряд сложностей при работе с камерой и с использованием самого робота без поддержки внешнего ПК как основного узла для расчетов. Поэтому все построения на данном этапе производились при помощи программной эмуляции на компьютере.

Не смотря на возникшие сложности и проблемы, удалось реализовать комплекс, который может быть сравнительно легко перенесен на другую физическую платформу и может быть использован при создании более сложных интеллектуальных систем.

Результаты исследований показали успешность разработанного метода и наметили пути дальнейшего развития.

Литература

1. Притыкин Ф.Н., Небритов В. И. Определение максимальных значений параметров задающих угол сервиса и множество конфигураций андроидного робота реализацией мгновенных состояний // Инженерный вестник Дона, 2016. № 1. URL: ivdon.ru/ru/magazine/archive/n1y2016/3506/.

2. Аджиева А.А. Шаповалов В.А. Кластерный анализ в автоматическом выявлении и сопровождении грозовых очагов по данным грозопеленгационной сети // Инженерный вестник Дона, 2016. № 2. URL: ivdon.ru/ru/magazine/archive/n2y2016/3559/.

3. Гинис Л.А. Развитие инструментария когнитивного моделирования для исследования сложных систем // Инженерный вестник Дона, 2013. № 3. URL: ivdon.ru/ru/magazine/archive/n3y2013/1806/.
4. Крашенинников А.М., Гданский Н.И., Рысин М.Л. Построение сложных классификаторов для объектов в многомерных пространствах // Инженерный вестник Дона, 2013. №2. URL: ivdon.ru/magazine/archive/n2y2013/1611/.
5. Singh, S., Kearns, M. Near-Optimal Reinforcement Learning in Polynomial Time. // Machine Learning journal. Vol. 49, Is. 2. 2002. pp.209-232.
6. Touzet C. Q-learning for robots // The Handbook of Brain Theory and Neural Networks: second edition / M. A. Arbib ed., MIT Press, 2003. pp. 934-937.
7. Parker L. E., Touzet C., and Fernandez F. Techniques for Learning in Multi-Robot Teams // Robot Teams: From Diversity to Polymorphism / Tucker Balch and Lynne E. Parker eds. A K Peters/CRC Press, 2002. pp. 191-236.
8. Ziqiang P., Gang P., Ling Y. Learning Biped Locomotion Based on Q-Learning and Neural Networks // Advances in Automation and Robotics / Lee G. eds. Vol.1. (LNEE, vol.122). Berlin, Heidelberg: Springer, 2011. pp. 313–321.
9. Mondada F., Bonani M., Raemy X., Pugh J., Cianci C., Klaptocz A., Magnenat S., Zufferey J.-C., Floreano D. and Martinoli A. The e-puck, a robot designed for education in engineering // Proceedings of the 9th conference on autonomous robot systems and competitions, 2009. 1(1). pp. 59-65.
10. Sharkey N.E. The new wave in robot learning // Robotics and Autonomous Systems. 1997. V. 22. № 3-4. pp. 179-185.
11. Рассел С., Норвиг П. Искусственный интеллект: современный подход. 2-е издание. М.: Вильямс, 2007. 1408 с.

References

1. Pritykin F.N., Nebritov V. I. Inženernyj vestnik Dona (Rus), 2016. № 1. URL: ivdon.ru/ru/magazine/archive/n1y2016/3506/.



2. Adzhieva A.A. Shapovalov V.A. Inzhenernyj vestnik Dona (Rus), 2016. № 2. URL: ivdon.ru/ru/magazine/archive/n2y2016/3559/.
3. Ginis L.A. Inzhenernyj vestnik Dona (Rus), 2013. № 3. URL: ivdon.ru/ru/magazine/archive/n3y2013/1806/.
4. Krasheninnikov A.M., Gdanskiy N.I., Rysin M.L. Inzhenernyj vestnik Dona (Rus), 2013. №2. URL: ivdon.ru/magazine/archive/n2y2013/1611/.
5. Singh, S., Kearns, M. Machine Learning journal. Volume 49, Issue 2. 2002. pp. 209-232.
6. Touzet C. The Handbook of Brain Theory and Neural Networks: second edition. M. A. Arbib ed., MIT Press, 2003. pp. 934-937.
7. Parker L. E., Touzet C., and Fernandez F. Robot Teams: From Diversity to Polymorphism. Tucker Balch and Lynne E. Parker eds. A K Peters/CRC Press, 2002. pp. 191-236.
8. Ziqiang P., Gang P., Ling Y. Advances in Automation and Robotics. Lee G. eds. Vol.1. (LNEE, vol.122). Berlin, Heidelberg: Springer, 2011. pp. 313–321.
9. Mondada F., Bonani M., Raemy X., Pugh J., Cianci C., Klaptocz A., Magnenat S., Zufferey J.-C., Floreano D. and Martinoli A. Proceedings of the 9th conference on autonomous robot systems and competitions, 2009. 1(1). pp. 59-65.
10. Sharkey N.E. Robotics and Autonomous Systems. 1997. V. 22. № 3-4. pp. 179-185.
11. Rassel S., Norvig P. Искусственный интеллект: современный подход. [Artificial intelligence: a modern approach]. 2-е издание. М.: Vil'yams, 2007. 1408 p.