

Алгоритмы активного восприятия внешней среды для автономного мобильного робота

Д.В. Верхоланцев

МИРЭА – Российский технологический университет, г. Москва, Россия

Аннотация – Статья посвящена изучению алгоритмов активного восприятия для максимизации получаемой информации о среде функционирования автономного мобильного робота. Наиболее эффективным методом решения поставленной задачи можно считать применение автоматизированной системы управления положением видеокамеры. Для оценки полноты осмотра внешней среды бортовой видеокамерой использован метод конечных элементов. Применение простейшего интеллектуального алгоритма выбора направления осмотра в двухмерной симуляции позволило добиться 12% прироста эффективности для помещений из тестового набора.

Ключевые слова – робототехника, алгоритмы, активное восприятие, мобильные роботы.

I. ВВЕДЕНИЕ

Множество компаний по всему миру ведут активные исследования и разработки в области создания сервисных роботов для решения разнообразных задач, среди которых сопровождение и помощь людям, доставка грузов, патрулирование территории, выполнение работ на складах и заводах и т.д. На первый план при создании таких роботов выходит потребность в развитии технологий технического зрения, обеспечении скорости и точности детекции объектов, достижении полноты формируемой карты внешней среды и повышении степени интерактивности взаимодействия с пользователем.

В связи с повышением потребности в роботизации предлагаются различные варианты решения вышеупомянутых задач. Ведутся активные разработки в области алгоритмов распознавания и трекинга объектов, картографирования и локализации с применением подходов на основе нейронных сетей. Параллельно быстрыми темпами развиваются вычислительные устройства, ориентированные на ускорение обработки нейросетевых моделей. Однако, стоит заметить, что данных улучшений недостаточно, так как их результативность напрямую зависит от получаемой информации. Здесь на передний план выходит усовершенствование алгоритмов восприятия.

Одной из первых работ в этой области принято считать диссертацию Дж. Тененбаума, посвящённую

аккомодации в компьютерном зрении [1]. Автор описывает подход отличный от обычных систем обработки изображений тем, что аккомодация сенсора происходит автоматически и рассматривается как неотъемлемая часть процесса распознавания. Под аккомодацией в указанной работе предполагается понимать автоматическое изменение фокусного расстояния и управление фильтром нейтральной плотности для изменения яркости изображения.

Иную методику разрешения неопределённости в своей работе предлагают Х. Барроу и Р. Попплетоун [2]. Исследователи указывают на возможность получения большего количества информации путём изменения точки обзора или положения исследуемого объекта.

Определение концепции активного восприятия появляется в работе Дж. Алаимоноса и других исследователей, посвящённой активному зрению [3]. Учёные предлагают следующую концепцию: “Наблюдатель называется активным, когда он занят какой-либо деятельностью, целью которой является контроль геометрических параметров сенсорного аппарата. Целью деятельности является манипулирование ограничениями, лежащими в основе наблюдаемых явлений, с целью улучшения качества результатов восприятия”.

Эти концепции нашли своё применение в решении многих задач в области робототехники. Отдельным направлением активного восприятия можно считать системы активной одновременной локализации и картографирования (Active SLAM, ASLAM), которым посвящена соответствующая работа [4]. Эти исследования концентрируются на активном расчёте траекторий перемещения мобильных роботов для получения карты местности с минимальной ошибкой.

Иное применение системы активного восприятия нашли в сельском хозяйстве для контроля и сбора урожая [5]. Здесь ведутся активные исследования, направленные на создание алгоритма управления моторизованным взглядом.

Альтернативный вариант понимания активного восприятия в своей работе [6] представили исследователи из университета Васэда. Они указывают на необходимость активного построения взаимосвязей инструмент – объект – действие при взаимодействии агента с окружающей средой.

Также на потребность в активном исследовании указывают авторы работы, посвящённой визуально-языковой навигации [7]. Здесь система целенаправленного изучения окружающего пространства позволяет устранять неопределённости при выполнении поставленных задач.

II. ПОСТАНОВКА ЗАДАЧИ

Важным условием эффективности выполнения задачи робототехнической системой является достаточность объёма информации о среде функционирования. В случае же решения задачи мониторинга внешней среды максимизация информации об области изучения является приоритетным процессом. Отсюда возникает потребность в создании подсистемы, осуществляющей целенаправленные действия, обеспечивающие увеличение доступной для обработки информации.

Пусть положение робота определяется вектором линейных координат и угла ориентации на плоскости $p_r = \{x_r, y_r, \alpha_r\}$, а маршрут робота задан множеством точек $T = \{p_1, \dots, p_N\}$. Положение видеокамеры задается вектором $p_c = \{x_c, y_c, \alpha_c\}$, где угол поворота α_c определяется согласно выбранному закону. Дискрета местности имеет вид $d = \{x, y, v\}$, где v – отметка посещённости. Число дискрет местности составляет $n_x * n_y$, $D = \{d_1, \dots, d_N\}$ – множество всех дискрет, $D^* = \{d_{i1}, \dots, d_{ik}\}$ – множество осмотренных дискрет.

Пусть дополнительно на множестве D задан граф местности $G = \{V, E\}$, где V – множество вершин, E – множество ребер. Причем местоположения вершин совпадают с положением дискрет местности. Граф G может быть использован для поиска первичного (опорного) маршрута движения робота при осмотре внешней среды.

Требуется найти такой закон изменения угла видеокамеры α_c , при котором выполняется $|D^*| \rightarrow \max$.

III. МОДЕЛИ И АЛГОРИТМЫ АКТИВНОГО ВОСПРИЯТИЯ

А. Определение понятия активного восприятия в задачах сервисной робототехники

Восприятие – процесс обработки данных от сенсорного оснащения робота с целью выявления информации необходимой для последующего выполнения поставленной перед робототехнической системой задачи. Стандартное восприятие происходит параллельно основной деятельности сервисного робота и не предполагает целенаправленного осуществления действий, направленных исключительно на расширение имеющейся информации о среде функционирования.

Стандартный вариант реализации системы восприятия подразумевает ограничение получаемой информация следующим набором факторов:

- физическими параметрами сенсоров (угол обзора и фокусное расстояние видеокамеры, система развёртки и дальность действия лидара);
- конструкцией робота (наличие видеокамеры спереди не позволяет получать информацию об объектах, расположенных сзади);
- имеющейся вычислительной мощностью (увеличение числа сенсоров приведёт к увеличению количества ресурсов, требуемых для обработки их данных);
- параметрами сцены (необходимый объект может быть перегорожен другим объектом сцены).

При этом действия робота направлены на достижение поставленной цели на основе имеющейся информации, но не на её расширение. Таким образом возникает вероятность недостижимости поставленной цели из-за неполноты доступной ему информации.

Анализ существующих трактовок [1–3] позволяет сделать обобщение и понимать активное восприятие как процесс осуществления роботом действий, направленных на расширение объёма получаемых данных путём изменения физических параметров сенсоров, собственной конструкции или параметров сцены на основе уже имеющейся информации о среде функционирования и поставленной задачи.

Подход, основанный на использовании системы активного восприятия, позволяет получать более полную и актуальную информацию о среде функционирования робота, что в свою очередь позволяет с большей результативностью выполнять поставленные задачи. Многочисленные исследования предлагают множество методов решения описанной проблемы. Однако, далеко не все подходы к реализации таких систем в равной степени эффективны.

Автоматическое изменение физических параметров сенсоров (фокусного расстояния, угла обзора, дальности функционирования дальнометрических сенсоров, плотности облака точек лидара и др.) зачастую является невозможным при использовании стандартных аппаратных решений, а разработка проприетарных устройств приведёт к удорожанию итоговой системы и снизит её надёжность, усложнив конструкцию и, как следствие, увеличив вероятность ошибки. Изменение параметров сцены так же может быть невозможным из-за её особенностей, таких как неподвижность определённых объектов в ней. Таким образом, этот метод сводится к изменению точки наблюдения путём смещения сенсорного блока (рис. 1).

Добиться этого можно несколькими способами. Первый подразумевает статичную конфигурацию сенсорного оборудования и предполагает перемещение всего робота.

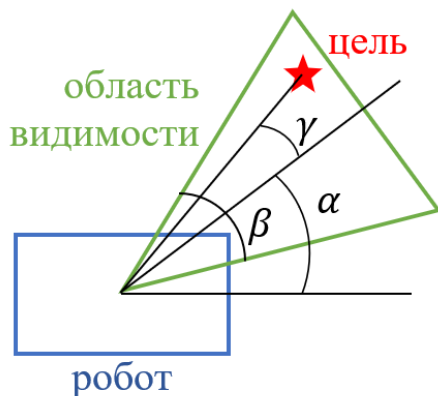


Рис. 1. Геометрическая схема системы активного восприятия на базе поворотной бортовой видеокамеры

Недостатком такого подхода является необходимость перестройки маршрута движения робототехнической системы, осуществление которого потребует дополнительных временных и энергетических затрат. При этом существенные ограничения на систему накладывают конструкционные параметры робота, небольшие размеры не позволят наблюдать объекты, расположенные, например, на столе, а увеличение габаритов приведёт к ухудшению проходимости и сделает невозможным достижение труднодоступных мест. Избежать вышеуказанных проблем позволит конструкция робототехнической системы с автоматически изменяемой конфигурацией сенсорного оснащения. Именно этот подход будет использоваться в дальнейшей работе.

В. Алгоритмы активного восприятия

Поскольку мобильный робот сочетает в себе подвижную платформу и поворотную видеокамеру, то алгоритм активного восприятия, так или иначе, опирается на управление движением обеих подсистем. На уровне подвижной платформы маршрут робота планируется исходя из необходимости посещения всех вершин графа местности. Здесь можно заметить, что поставленную задачу можно рассматривать как частный случай задачи коммивояжёра.

Одним из вариантов решения подобных задач является метод поиска в глубину (Depth-first search, DFS) [8]. Основная идея алгоритма заключается в рекурсивном исследовании ранее не исследованных вершин. Случайный выбор вершин для исследования делает результирующую траекторию, то есть последовательность узлов графа, неоптимальной. Выбор перехода на основании минимальности веса ребра делает исходный алгоритм жадным, то есть локально оптимальным для каждого шага. Необходимо отметить, что в подавляющем большинстве случаев результат работы жадного алгоритма не будет оптимален глобально.

Для оптимизации такого рода задач предлагается развитый набор эволюционных алгоритмов [9]. В этой

же работе предлагается сконцентрироваться на управлении направлением взгляда.

Использование автоматической системы поворота видеокамеры даёт возможность сокращать опорный маршрут робота за счет заблаговременного осмотра точек местности. Рассмотрим несколько сценариев поведения бортовой видеокамеры. Может показаться логичным использовать бортовую видеокамеру для постоянного сканирования окружения. Отсюда первый вариант реализации системы активного восприятия предполагает равномерное постоянное вращение бортовой видеокамеры в плоскости сцены. Таким образом, закон управления взглядом можно формально описать следующим образом:

$$\frac{\Delta \alpha_c}{\Delta t} = \text{const} \quad (1)$$

Здесь необходимо отметить, что слишком высокая скорость вращения приведёт к снижению качества видеопотока и, следовательно, негативно скажется на качестве работы алгоритмов обработки изображения. Слишком низкая же скорость сделает систему нецелесообразной. Эмпирически была выбрана скорость вращения, равная 0,5 рад/сек. Также нужно заметить, что реализация подобной конструкции при использовании стандартных для индустрии видеокамер представляется затруднительным.

Другой возможный подход предполагает использование интеллектуального алгоритма поиска наиболее результативного направления осмотра бортовой видеокамеры. Как упоминалось ранее для определения осмотренной территории предполагается использовать дискретизированное представление доступного пространства, где каждая дискрета из множества всех дискрет D может также попадать в подмножество исследованных дискрет D^* . Направление осмотра задаётся последовательностью из доступных опорных направлений, представляющих собой набор $A = \{\alpha'_1, \dots, \alpha'_{10}\}$ из десяти значений в диапазоне от -170 до 170 градусов с шагом в 34 градуса. Выбор целевого направления взгляда описывается:

$$\alpha_s = \underset{\alpha'}{\operatorname{argmin}} |D^* \cap F(\alpha')| \quad (2)$$

где, α_s – опорное направление взгляда на текущем этапе функционирования системы;

$F(\alpha')$ – множество дискрет, попадающих в область видимости бортовой видеокамеры при выбранном направлении.

Скорость перемещения между опорными направлениями, которые являются результатом работы алгоритма, останется неизменной и будет равна 0,5 рад/сек. Таким образом, итоговый закон движения бортовой видеокамеры описывается выражением:

$$\alpha_c^t = \alpha_c^{t-1} + \frac{\alpha_s^t - \alpha_c^{t-1}}{|\alpha_s^t - \alpha_c^{t-1}|} * \frac{0.5 * 180}{\pi} \quad (3)$$

В качестве контрольного метода предлагается использовать классический вариант с фиксированным относительно мобильной платформы положением бортовой видеокамеры. Закон управления в таком случае описывается:

$$\alpha_c = const \quad (4)$$

IV. СИМУЛЯЦИЯ СТРАТЕГИЙ АКТИВНОГО ВОСПРИЯТИЯ

Как упоминалось ранее, основной целью подсистемы активного восприятия является максимизация осмотренной в процессе функционирования площади. Именно этот аспект будет исследоваться в процессе симуляции. В качестве навигационной задачи для мобильной платформы было выбрано перемещение по траектории, заданной последовательностью опорных точек. Для ускорения тестирования воспользуемся упрощением в виде двумерной проекции симулируемой сцены (рис. 2).

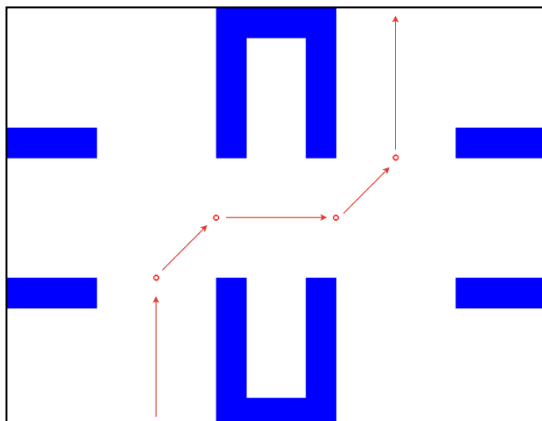


Рис. 2. Схема симулируемой сцены, где синим обозначены стены, а красным направление движения и опорные точки траектории

Для упрощения проводимых исследований предлагается сделать допущение о дискретизации свободного от препятствий пространства сцены. Для определения эффективности осмотра территории будет использоваться следующая формула:

$$Q = \frac{N_1}{N_2} * 100\% \quad (5)$$

где N_1 — число дискрет местности, которые попали в область видимости, принадлежащий бортовой видеокамере робота;

N_2 — общее число дискрет местности.

V. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ

Использование статичной (рис. 3) видеокамеры позволило исследовать 433 области из 864, то есть $Q \approx 50\%$. Вращающаяся видеокамера (рис. 4) позволила исследовать 456 областей из 864, то есть $Q \approx 53\%$. Применение интеллектуального алгоритма управления видеокамерой (рис. 5) позволило исследовать 538 областей из 864, то есть результативность осмотра составила 62% при неизменных прочих параметрах.

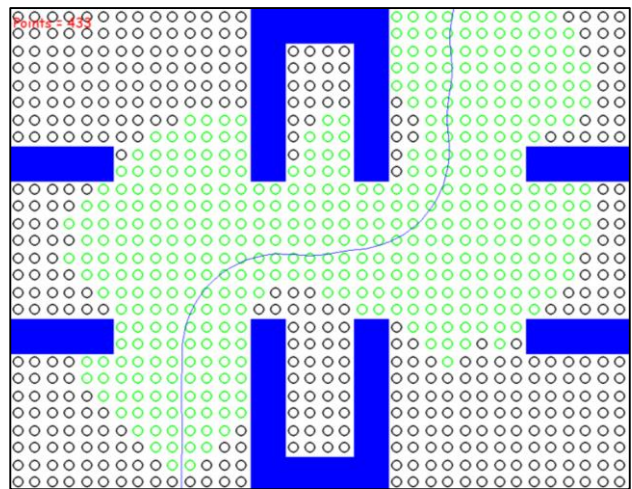


Рис. 3. Результат симуляции со статичной видеокамерой

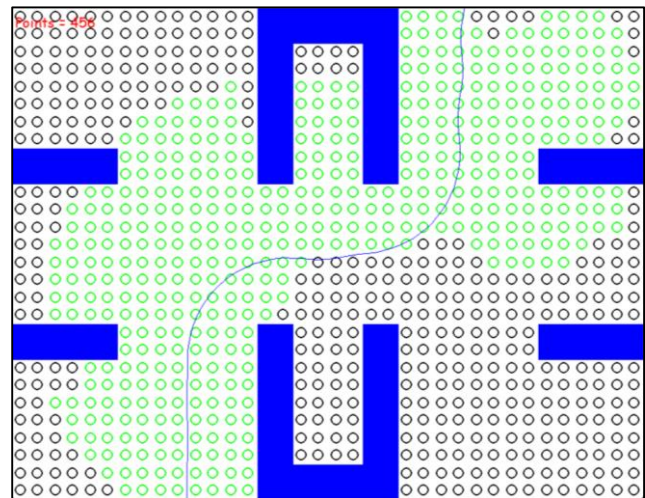


Рис. 4. Результат симуляции с равномерно вращающейся видеокамерой

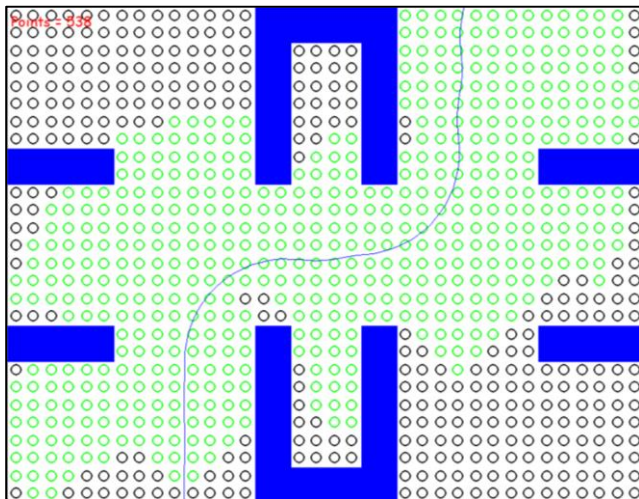


Рис. 5. Результат симуляции с интеллектуальным алгоритмом управления видеокамерой

VI. ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Анализ результатов симуляции вышеупомянутых алгоритмов позволяет сделать вывод о превосходстве алгоритмов активного восприятия над пассивными вариантами реализации подобных систем. В задачах максимизации осматриваемой площади результат применения интеллектуального алгоритма осмотра позволил добиться прироста в 12% эффективности, в то время как равномерно вращающаяся видеокамера позволила исследовать лишь на 3% больше территории.

VII. ВЫВОДЫ И ЗАКЛЮЧЕНИЕ

В работе был проведён анализ существующих исследований в области активного восприятия для решения различных задач, поставленных перед робототехническими системами. В контексте задачи мониторинга местности для мобильной платформы был сделан вывод о необходимости использования активного восприятия как автоматической системы управления направлением взгляда бортовой видеокамеры для максимизации осмотренной площади без необходимости дополнительных изменений в траектории перемещения самой мобильной платформы.

Такой подход позволяет сократить временные и энергетические затраты на выполнение поставленной задачи. Результаты, полученные в ходе симуляции, демонстрируют, что простое равномерное вращение бортовой видеокамеры позволяет добиться трёхпроцентного прироста в эффективности, что при упомянутых ранее недостатках такой конструкции можно считать несущественным улучшением.

Применение же интеллектуального алгоритма управления направлением бортовой видеокамеры, основанного на анализе попадающей в область видимости ранее неисследованной площади, с в

дальнейшем планируется перейти к трёхмерной модели окружающего пространства и провести испытания итоговой системы на реальном оборудовании.

ЛИТЕРАТУРА

- [1] Tenenbaum J. M. Accommodation in computer vision // Ph.D. Thesis, Stanford University. 1970.
- [2] Barrow H. G., Poppelstone R. J. Relational descriptions in picture processing // Machine Intelligence. 1971, V. 6. P. 377–396.
- [3] Aloimonos J., Weiss I., Bandyopadhyay A. Active vision // International journal of computer vision. 1988, V. 1. C. 333–356.
- [4] Lluvia I., Lazkano E., Ansuategi A. Active mapping and robot exploration: A survey // Sensors. 2021. T. 21. №. 7. C. 2445.
- [5] Magalhães, S. A., Moreira, A. P., Santos, F. N. D., Dias, J. Active perception fruit harvesting robots—A systematic review // Journal of Intelligent & Robotic Systems. 2022. T. 105. №. 1. C. 14.
- [6] Saito, N., Ogata, T., Funabashi, S., Mori, H., Sugano, S. How to select and use tools?: Active perception of target objects using multimodal deep learning // IEEE Robotics and Automation Letters. 2021. T. 6. №. 2. C. 2517–2524.
- [7] Wang, H., Wang, W., Liang, W., Hoi, S. C., Shen, J., Gool, L. V. Active perception for visual-language navigation // International Journal of Computer Vision. 2023. T. 131. №. 3. C. 607–625.
- [8] Nasirian B., Mehrandezh M., Janabi-Sharifi F. Efficient coverage path planning for mobile disinfecting robots using graph-based representation of environment // Frontiers in Robotics and AI. 2021. T. 8. C. 624333.
- [9] Tung W. C., Liu J. S. Solution of an integrated traveling salesman and coverage path planning problem by using a genetic algorithm with modified operators // IADIS Int. J. Comput. Sci. Inf. Syst. 2019. T. 14. №. 2. C. 95–114.

Информация об авторе:

Верхоланцев Дмитрий Витальевич, магистрант кафедры проблем управления Российского технологического университета - МИРЭА, г. Москва, Россия, e-mail: verkhold@gmail.com

Algorithms for active perception of the external environment for an autonomous mobile robot

D. V. Verkholtantsev

MIREA – Russian Technological University

Abstract – This article is devoted to the study of active perception algorithms for maximizing the information received about the operating environment of an autonomous mobile robot. The most effective method for solving the problem can be considered the use of an automated system for controlling the position of the video camera. The finite element method was used to assess the completeness of the inspection of the external environment by an onboard video camera. The use of the simplest intelligent algorithm for selecting the inspection direction in a two-dimensional simulation allowed us to achieve a 12% increase in efficiency for the premises from the test set.

Keywords – robotics, algorithms, active perception, mobile robots.

References

- [1] Tenenbaum J. M. Accommodation in computer vision // Ph.D. Thesis, Stanford University. 1970.
- [2] Barrow H. G., Poppelstone R. J. Relational descriptions in picture processing // Machine Intelligence. 1971, V. 6. P. 377–396.
- [3] Aloimonos J., Weiss I., Bandyopadhyay A. Active vision // International journal of computer vision. 1988, V. 1. C. 333–356.
- [4] Lluvia I., Lazkano E., Ansuategi A. Active mapping and robot exploration: A survey // Sensors. 2021. T. 21. №. 7. C. 2445.
- [5] Magalhães, S. A., Moreira, A. P., Santos, F. N. D., Dias, J. Active perception fruit harvesting robots—A systematic review // Journal of Intelligent & Robotic Systems. 2022. T. 105. №. 1. C. 14.
- [6] Saito, N., Ogata, T., Funabashi, S., Mori, H., Sugano, S. How to select and use tools?: Active perception of target objects using multimodal deep learning // IEEE Robotics and Automation Letters. 2021. T. 6. №. 2. C. 2517–2524.
- [7] Wang, H., Wang, W., Liang, W., Hoi, S. C., Shen, J., Gool, L. V. Active perception for visual-language navigation // International Journal of Computer Vision. 2023. T. 131. №. 3. C. 607–625.
- [8] Nasirian B., Mehrandezh M., Janabi-Sharifi F. Efficient coverage path planning for mobile disinfecting robots using graph-based representation of environment // Frontiers in Robotics and AI. 2021. T. 8. C. 624333.
- [9] Tung W. C., Liu J. S. Solution of an integrated traveling salesman and coverage path planning problem by using a genetic algorithm with modified operators // IADIS Int. J. Comput. Sci. Inf. Syst. 2019. T. 14. №. 2. C. 95–114.