

УДК 621.865 : 004.896 : 623

МЕТОДЫ ВИЗУАЛЬНОЙ НАВИГАЦИИ МОБИЛЬНОГО РОБОТА И ПОСТРОЕНИЯ КАРТОГРАФИЧЕСКИХ МОДЕЛЕЙ ВНЕШНЕЙ СРЕДЫ

© Ж. С. Першина¹, С. Я. Каздорф¹, А. В. Лопота²

¹Новосибирский государственный технический университет,
630070, г. Новосибирск, просп. К. Маркса, 20

²Центральный научно-исследовательский и опытно-конструкторский институт
робототехники и технической кибернетики,
194064, Санкт-Петербург, Тухоревский пр., 21
E-mail: pershina@tiger.cs.nstu.ru
alopota@rtc.ru

Рассмотрены современные методы визуальной навигации мобильного робота. Предложена иерархическая структура представления окружающего пространства, соответствующая иерархической организации системы управления мобильного робота. Представлены современные подходы к построению картографических моделей. Их развитие позволит приблизить систему навигации к той, которая реализуется интеллектом человека и объединяет зрение и семантическое представление о мире в рамках когнитивных карт.

Ключевые слова: визуальная навигация, свёрточные нейронные сети, семантическая сегментация, когнитивная карта.

DOI: 10.15372/AUT20190210

Введение. Одним из актуальных направлений развития наземной мобильной робототехники является совершенствование управления мобильными роботами, включая навигацию в условиях неопределённой и динамически меняющейся внешней среды. В настоящее время известен ряд наземных мобильных роботов, обладающих относительной автономностью при решении задач перемещения в соответствии с выполняемыми задачами. Как правило, их самостоятельность ограничена возможностью планирования оптимальной и безопасной траектории на основе метрической информации о наблюдаемой сцене. Высокий уровень автономности может быть достигнут лишь за счёт полного использования всех каналов информации об окружающем мире при решении задач навигации, включая семантический уровень представления данных [1, 2].

Перспективный способ такого управления основан на использовании визуальной информации о внешней среде. Зрение — один из основных и древних в эволюционном смысле способов восприятия информации, имеющий большую интуитивную привлекательность, поскольку люди и животные при навигации в пространстве зачастую пользуются именно зрительной информацией. В процессе восприятия зрительной информации живой организм оперирует некими целостными образами среды или системой образов, последовательно их обобщая в процессе обработки и выявляя пространственно-временные отношения между ними [1]. Задача обеспечения навигации мобильных роботов та же, что и у живых существ в природе. Поэтому естественно рассмотреть перспективность такого нового подхода и в робототехнике.

Подход, основанный на использовании визуальной информации при навигации мобильного робота, получил название «Визуальная навигация». Задачами навигации мобильного робота являются:

- картографирование внешней среды,
- оценка текущего местоположения мобильного робота,
- планирование траектории движения.

Цель данного обзора — рассмотреть современное состояние решения задач картографирования внешней среды на основе визуальной информации, получаемой от камеры, которая закреплена на борту мобильного робота, и одновременно определить на формируемой в процессе движения карте его текущее местоположение. В зарубежных исследованиях данное направление получило название vSLAM (visual Simultaneous Localization and Mapping) [3–7] и в некоторых работах известно как TAM (Tracking and Mapping) [8, 9].

Классификация способов визуальной навигации. Для визуальной навигации применяются три типа камер: монокулярная, стерео и RGB-D. В отличие от монокулярных камер стереокамеры и камеры RGB-D (Red, Green, Blue, Depth) позволяют получить карту глубины (depth map) поверхностей объектов, попавших в поле зрения камеры. Использование камер для осуществления навигации мобильного робота обосновано тем, что они обладают широкой полосой пропускания по сравнению с другими сенсорными системами и соответственно наибольшими семантическими возможностями для распознавания образов при картографировании, тем самым предоставляя возможность осуществлять движение по внешним ориентирам. Последнее получило название «Интерпретирующая навигация» [10]. Распознанные с помощью видеокамер визуальные ориентиры позволяют путём обнаружения их на карте также устанавливать истинные координаты и планировать движение не только по метрической карте, но и по топологическому образу среды.

При этом визуальные ориентиры можно классифицировать как искусственно созданные и естественные. Выделение и отслеживание искусственных ориентиров упрощает задачу, так как они разрабатываются с оптимальным контрастом и заранее известны их точные геометрические и физические свойства. Первыми искусственными ориентирами, использовавшимися в робототехнике, являлись контрастные по яркости с фоном линии на полу и их пересечения. Рассогласование между текущим и целевым изображениями линии, вдоль которой должен двигаться робот, поступало на вход системы управления роботом, таким образом позволяя корректировать траекторию движения. Данный подход в системах управления известен как «Визуальное сервоуправление» [11, 12].

К другим, наиболее удобным для распознавания, искусственным ориентирам относятся чёрно-белые объекты-эталонные простой формы. Объекты простой формы могут формироваться, например, на основе технологии двумерного штрихового кодирования [13–14].

В работе [15] приведён пример навигации группы мобильных роботов по искусственным ориентирам — кубическим объектам с нанесёнными на их стороны QR-кодами. Каждый робот группы осуществляет поиск ориентира, вычисляет его положение относительно собственной стартовой позиции и выполняет построение локальной карты. Параллельно этому локальные карты от различных роботов интегрируются в глобальную карту и уточняются с течением времени.

При решении навигационных задач в неструктурированных внешних средах, в которых использование искусственных, заведомо определённых ориентиров невозможно, применяется подход к выделению естественных визуальных ориентиров.

На сегодняшний день данный подход можно классифицировать по методам анализа изображений:

- 1) вычисление дескрипторов локальных окрестностей особых точек на изображении;
- 2) попиксельный анализ изображения и карты глубины;
- 3) семантический анализ изображения.

При использовании в качестве визуальных ориентиров особых точек (point features), которые являются инвариантными к изменению масштаба, поворота и освещения, оценка текущего местоположения мобильного робота осуществляется на основе сопоставле-

ния и определения смещения особых точек между полученными кадрами изображений. Под оценкой текущего местоположения понимается оценка смещения мобильного робота в трёхмерном пространстве, включая углы Эйлера (6DoF — Six Degrees of Freedom).

Общий алгоритм вычисления особых точек представлен последовательностью процедур:

- выбор метода и вычисление дескриптора локальных окрестностей особых точек;
- сопоставление особых точек;
- отсеечение ложных срабатываний.

Для описания локальных окрестностей особых точек на изображении зачастую используют такие дескрипторы, как SURF [16], SIFT [17], BRIEF [18], ORB [19] и др. Одним из быстрых и надёжных алгоритмов сопоставления особых точек является алгоритм «ближайшего соседа» (Nearest Neighbor Search, NNS), который по сравнению с методами простого перебора позволяет увеличивать производительность.

После сопоставления особых точек на двух соседних кадрах для отсеечения ложных срабатываний применяется алгоритм RANSAC (RANdom SAmple Consensus) [20], который выполняет оценку параметров модели на основании случайных выборок. При сопоставлении модель представляет собой матрицу трансформации (гомография). На входе алгоритма RANSAC имеются два множества дескрипторов, полученных на предыдущем и текущем кадрах. В результате применения RANSAC строится наилучшая матрица трансформации на основе формирования гипотез и определяется пороговое значение. В том случае если матрица трансформации пары дескрипторов текущего и предыдущего кадров не соответствует полученной матрице с учётом порогового значения, то пара отбрасывается.

Ключевыми алгоритмами в области визуальной навигации на основе вычисления дескриптора локальных окрестностей особых точек являются MonoSLAM [3], PTAM [8], ORB-SLAM [4].

MonoSLAM позволяет оценить 6DoF мобильного робота при сопоставлении и оценке смещения особых точек в каждом кадре изображения, полученного от монокулярной камеры с использованием расширенного фильтра Калмана (Extended Kalman Filter, EKF). Трёхмерные координаты особых точек и 6DoF мобильного робота представлены в виде вектора состояния в EKF, а в качестве модели движения используется равномерное движение. В процессе движения мобильного робота к вектору состояния добавляются координаты новых особых точек. Стоит отметить, что инициализация начального положения мобильного робота выполняется на основе вычисления особых точек известного объекта с функцией маркера, в системе координат которого определена глобальная система координат.

Главными недостатками данного метода являются накопление ошибки и увеличение вычислительной сложности. В больших средах появление новых особых точек ведёт к увеличению размера вектора состояния, что, в свою очередь, делает невозможным выполнение вычислений в реальном времени.

Для решения проблемы вычислительной сложности разработчики PTAM [8] предложили разделить отслеживание визуальных ориентиров и картографирование в разные вычислительные потоки. Эти два потока выполняются параллельно на разных потоках многоядерного процессора, так что вычислительная сложность картографирования не влияет на вычислительную сложность отслеживания, что позволяет использовать алгоритм в реальном масштабе времени.

Результатом дальнейшего развития методов vSLAM, базирующихся на использовании в качестве визуальных ориентиров особых точек, является алгоритм ORB-SLAM, в котором в качестве дескриптора выступает ORB (Oriented FAST and Rotated BRIEF) [19], обладающий высоким быстродействием и высокой степенью инвариантности к углу обзора, повороту камеры и освещённости. Это позволяет алгоритму с высокой точностью

и надёжностью выполнять картографирование и отслеживать текущее местоположение мобильного робота.

Методы, основанные на использовании в качестве визуальных ориентиров особых точек, отличаются достаточно высоким быстродействием, но несмотря на это имеют существенное ограничение: их невозможно применять в средах с отсутствием ярко выраженной текстуры.

К другому классу относятся методы, созданные на анализе изображений и карт глубины целиком — без извлечения абстрактных признаков: DTAM [9], DVO [21], LSD-SLAM [5].

В методе DTAM (Dense Tracking and Mapping) применяется последовательность RGB-изображений для реконструкции сцены и оценки текущего местоположения камеры. Отличительной особенностью данного алгоритма является использование каждого пикселя обрабатываемого изображения, что позволяет добиться большей точности в сравнении с методами, разработанными на особых точках.

Метод DVO (Dense Visual Odometry) основан на вычислении матрицы трансформации между двумя соседними кадрами RGB-D с учётом схожести значений цвета и глубины спроектированных 3D-точек. При этом определение матрицы трансформации каждого кадра сравнительно с предыдущим приводит к накоплению погрешности, вызванной шумом сенсоров и неточностью применяемой модели ошибки. Данную проблему минимизируют с помощью выбора ключевых кадров (keyframe), относительно которых определяется матрица трансформации.

Построение трёхмерной карты и установление позиции камеры в методе LSD-SLAM (Large-Scale Direct Monocular SLAM) также базируется на выборе ключевых кадров и оценке карты глубины из последовательности RGB-изображений. Ключевой кадр находится из оценки удалённости и отклонения текущего кадра от предыдущего ключевого кадра. Значения карты глубины оцениваются только в областях изображения с наибольшим контрастом, что позволяет снизить вычислительную сложность. Значения пикселей карты глубины для каждого ключевого кадра нормируются в диапазоне от 0 до 1. Для задания масштаба значений карты глубины каждого ключевого кадра при построении трёхмерной карты применяется алгоритм обработки, с помощью которого достраивается карта трёхмерного пространства.

В последнее время также выделился новый класс методов Semantic SLAM, объединяющий в том или ином виде элементы вышеописанных методов, и метод, основанный на распознавании объектов, участвующих в навигации [6, 7]. Такой подход, помимо картографирования и локализации мобильного робота в пространстве, предполагает локализацию конкретных объектов в данном пространстве с возможностью одновременного отображения их на карте, формируемой в процессе перемещения камеры.

Появление и развитие этого подхода определено последними достижениями в области глубокого машинного обучения (deep learning), основным направлением развития которого являются совершенствование методов распознавания образов с помощью свёрточных нейронных сетей.

Типы построения картографических моделей. Рассмотрим существующие подходы к построению картографических моделей окружающего пространства применительно к уровням иерархической организации системы управления мобильным роботом, в которой выделяются следующие уровни управления: интеллектуальный, стратегический, тактический и исполнительный [22].

Высшим уровнем управления в системе является интеллектуальный. На данном уровне планируется поведение мобильного робота в условиях неполной информации о внешней среде. Функции интеллектуального уровня в робототехнических системах обычно выполняет человек-оператор, однако актуальным направлением развития мобильной ро-

бототехники является совершенствование управления мобильными роботами в условиях неопределённой и динамически меняющейся внешней среды без участия человека.

Стратегический уровень управления предназначен для планирования движений мобильного робота. Планирование движений предполагает разбиение задачи движения, поставленной интеллектуальным уровнем, на последовательность согласованных во времени элементарных действий и формализацию целей управления для каждого из этих действий.

Формализация целей управления означает, что для каждого из элементарных действий должны быть записаны математические соотношения, выполнение которых обеспечивает успешную реализацию действия. Стратегический уровень выдаёт информацию о плане движения и целях управления в форме команд управления движением.

Тактический уровень выполняет преобразование команд управления движением, поступающих со стратегического уровня управления, в программу управления, которая определяет законы согласованного движения во времени всех звеньев механического устройства мобильного робота с учётом технических характеристик блока приводов (в первую очередь ограничений на обобщённые скорости, ускорения и силы).

Исполнительный уровень управления предназначен для расчёта и выдачи управляющих сигналов на блок приводов мехатронной системы в соответствии с программой управления с учётом технических характеристик силовых преобразователей.

Реализация и обеспечение автономной навигации мобильного робота требует такого же иерархического представления картографической модели окружающего пространства.

Так, на более высоких уровнях системы управления, а именно интеллектуальном и стратегическом, планирование поведения может выполняться при описании наблюдаемой сцены в терминах когнитивных карт.

Когнитивная карта — образ знакомого пространственного окружения. Термин предложен в 1948 г. американским психологом Э. Толменом [23]. Наблюдая за перемещением крыс по лабиринту, учёный пришёл к выводу, что в процессе обучения животные не обязательно прибегают к методу проб и ошибок, а способны составлять и запоминать некую образную картину лабиринта — когнитивную карту.

В рамках разработки системы навигации мобильного робота описание внешней среды в виде визуальных образов предполагает использование топологического подхода, который представляет собой описание окружающего пространства в виде графа. При этом узлы графа соответствуют визуальным ориентирам (например, дверные проёмы), которые связаны дугами [24].

Применение данного подхода требует извлечения высокоуровневых обобщённых признаков на изображении наблюдаемой сцены, полученном от бортовой видеокамеры в целях распознавания и локализации в пространстве визуальных ориентиров.

Наибольшего успеха в области распознавания образов на изображении достигли решения, основанные на глубоких свёрточных нейронных сетях, в процессе обучения которых выполняется извлечение обобщённых признаков. В области распознавания можно выделить четыре класса задач [25]:

- классификация изображений,
- детектирование объектов на изображении,
- одноуровневая семантическая сегментация объектов на изображении,
- двухуровневая семантическая сегментация объектов на изображении.

Задача классификации предполагает определение класса объектов, к которому относится изображение, при этом не решается задача указания положения объектов на изображении. К базовым архитектурам свёрточных нейронных сетей, предназначенных для классификации, относятся AlexNet [26], VGG [27], GoogLeNet [28], ResNet [29] и др.

При решении задач навигации данный подход может быть успешно применён при условии, что мобильному роботу необходимо распознать единственный объект на пред-

ставленном изображении либо определить тип наблюдаемой сцены. При этом для выполнения поиска и локализации объектов на изображении требуется использование либо алгоритмов детектирования, либо семантической сегментации.

Задача детектирования позволяет находить несколько объектов и определять их положение на одном изображении. Результатом детектирования является набор прямоугольных регионов на изображении, которые заключают в себе распознанные объекты соответствующих классов. Наиболее известными архитектурами свёрточных нейронных сетей в данной области являются Faster R-CNN [30], SSD [31], YOLO [32], которые успешно применяются при разработке систем навигации автономных мобильных роботов как внутри помещений, так и в условиях городской среды. Однако при разработке систем навигации, в которых требуются более точные алгоритмы планирования безопасной и оптимальной траектории движения мобильного робота, например при решении задач захвата и перемещения объектов, необходимо максимально точное определение контуров объектов на изображении. Данную задачу возможно решить с использованием семантической сегментации объектов на изображении.

Ключевыми архитектурами свёрточных нейронных сетей в области семантической сегментации объектов являются FCN [33], U-Net [34], SegNet [35], DeepMask [36]. Применение данного подхода обеспечивает попиксельную локализацию и классификацию объектов на изображении. При этом семантическую сегментацию можно разделить на два типа: одноуровневую и двухуровневую. Одноуровневая семантическая сегментация позволяет попиксельно локализовать объекты и классифицировать каждый пиксель, при этом не выполняет разделения экземпляров одного класса, в то время как двухуровневая решает эту проблему и помимо присваивания пикселю метки класса присваивает порядковый номер объекта. Метка класса даёт возможность определять весь класс, а порядковый номер — все объекты этого класса.

Обобщением образной информации является вербальное представление внешней среды, основанное на проблемно-ориентированных языках, близких к естественным. Оно включает описание знаний как об объектах, обуславливающих интерес к выполнению требуемой задачи, так и пространственно-временных отношениях между объектами и мобильным роботом на базе лингвистических переменных.

Такое изображение внешней среды предполагает преобразование количественных параметров в их качественное представление. Таким образом, лингвистические переменные позволяют интервально уточнять расстояние, взаимное положение, ориентацию по направлению движения, его скорость [1].

Одним из способов лингвистического представления внешнего мира с учётом пространственно-временных отношений является нечёткая логика. Лингвистические описания объектов внешней среды и отношений между ними задаются функциями принадлежности, которые обычно определяются экспериментально и учитывают особенности восприятия пространственных отношений человеком. Например, отношение удалённости может быть выражено лингвистической переменной «дистанция», термами которой являются следующие значения, определённые функциями принадлежности [37]: d_1 — очень близко, d_2 — близко, d_3 — не близко и не далеко, d_4 — далеко, d_5 — очень далеко.

Оценка ситуации для такой модели среды и принятие решений о действиях должны осуществляться на указанном проблемно-ориентированном языке с использованием фреймов и продукционных правил на основе нечёткой логики [1].

Реализация тактического и исполнительного уровней требует построения более точных картографических моделей на базе метрической информации.

В простейшем случае метрическая модель окружающего пространства может быть представлена в виде двумерной карты занятости (occupancy grid). Впервые формализованное описание этой модели предложено в [38]. Двумерная карта занятости являет собой

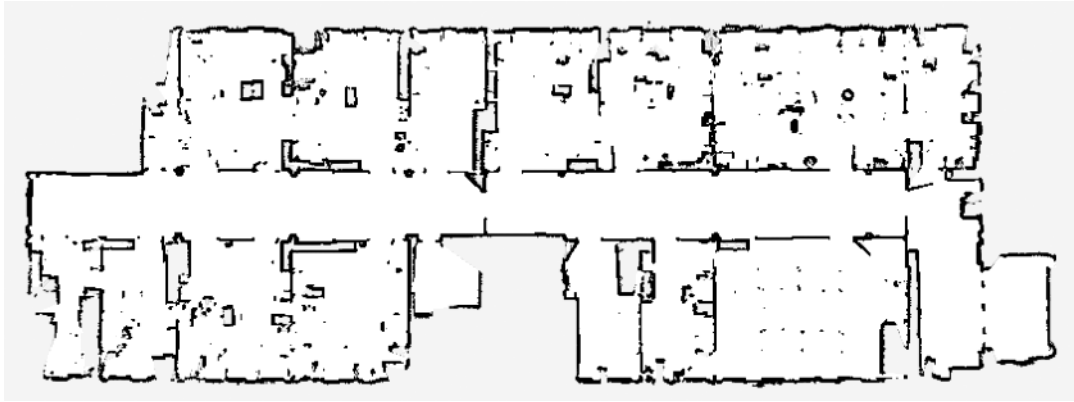


Рис. 1. Пример двумерной карты занятости помещения [38]

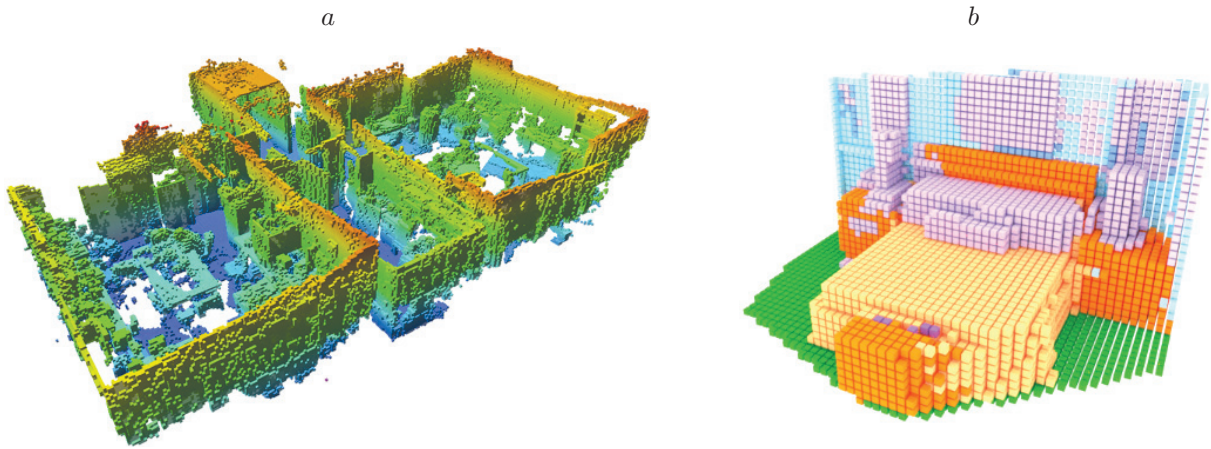


Рис. 2. Типы воксельных моделей: *a* — вероятностная модель занятости [39],
b — семантическая модель [40]

матрицу элементов, каждый из которых соответствует вероятности занятости определённого дискретного участка окружающего пространства. На рис. 1 дано отображение двумерной карты занятости [38].

Этот способ метрического описания окружающего пространства широко используется при решении задачи объезда препятствий и планирования траектории движения на свободных от препятствий участках. Однако такой способ не позволяет оценить геометрию рельефа и учесть высоту различных препятствий.

Другим способом метрического представления окружающего пространства является использование трёхмерной воксельной модели, основным элементом которой считается воксель (volume element, voxel). Значение каждого вокселя определяется либо вероятностью занятости [39] (рис. 2, *a*), либо значением семантической разметки, относящим принадлежность вокселя к конкретному классу объектов [40–42] (рис. 2, *b*).

Однако воксельные модели трудно применять при картографировании больших сред из-за высокой пространственной сложности и вычислительной стоимости при достаточной детализации. Это связано с тем, что разрешение метрических карт должно быть достаточно хорошим, чтобы захватывать важные элементы окружающего пространства.

Меньшей пространственной сложностью обладает модель в виде облака $3D$ -точек, представляющего собой набор вершин в трёхмерной системе координат. Эти вершины задаются координатами X , Y и Z и, как правило, предназначены для представления внешних поверхностей объектов (рис. 3, *a*).

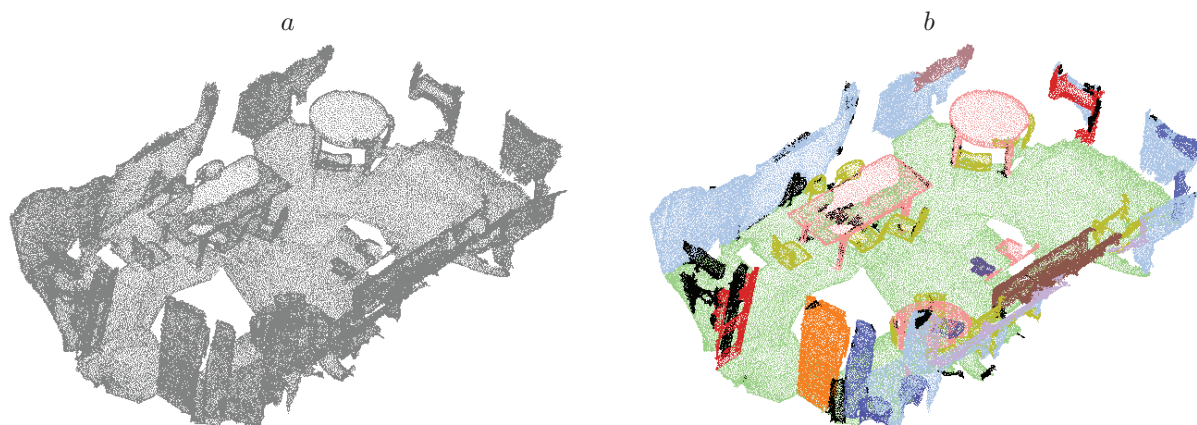


Рис. 3. Пример разметки облака 3D-точек: *a* — неразмеченное облако 3D-точек, *b* — семантически размеченное облако 3D-точек [43]

Дополнительная разметка облака 3D-точек позволяет путём проекции пиксельных координат, распознанных с использованием свёрточных нейронных сетей визуальных ориентиров на двумерном изображении, выполнить семантическую сегментацию объектов в трёхмерном пространстве (рис. 3, *b*) [44–48].

Заключение. В данной работе рассмотрены современные методы визуальной навигации мобильного робота. Существующие подходы к построению картографических моделей окружающего пространства сопоставляются с уровнями иерархической организации системы управления мобильного робота. На основании этого сделан вывод, что реализация и обеспечение автономной навигации мобильного робота требует такого же иерархического представления окружающего пространства. Этот подход даёт возможность приблизить систему навигации и определение ориентации робота к той, которая реализуется интеллектом человека и объединяет зрение и семантическое представление о мире. Дальнейшее развитие подхода семантической сегментации объектов в трёхмерном пространстве и формирование на его основе когнитивных карт окружающего пространства позволит обеспечить большую автономность и надёжность управления в неполнозаданных или априорно неопределённых условиях, что, в свою очередь, повысит безопасность и эффективность выполнения сложных операций.

Финансирование. Работа выполнена при поддержке Российского фонда фундаментальных исследований (проект № 18-58-76003).

СПИСОК ЛИТЕРАТУРЫ

1. **Половко С. А., Смирнова Е. Ю., Юревич Е. И.** Качественное управление мобильными роботами // Робототехника и техническая кибернетика. 2014. 4, № 3. С. 30–33.
2. **Лопота А. В., Половко С. А., Смирнова Е. Ю., Плавинский М. Н.** Основные результаты и перспективные направления исследований в области навигации и управления мобильными робототехническими комплексами // Исследования наукограда. 2013. 4, № 2. С. 49–53.
3. **Davison A. J., Reid I. D.** MonoSLAM: Real-time single camera SLAM // IEEE Trans. Pattern Analysis & Machine Intell. 2007. N 6. P. 1052–1067.
4. **Mur-Artal R., Montiel J. M. M., Tardos J. D.** ORB-SLAM: A versatile and accurate monocular SLAM system // IEEE Trans. Robotics. 2015. 31, N 5. P. 1147–1163.

5. **Engel J., Schöps T., Cremers D.** LSD-SLAM: Large-scale direct monocular SLAM // Proc. of the Eur. Conf. on Computer Vision (ECCV 2014). Zurich, Switzerland, 6–12 Sept., 2014. P. 834–849.
6. **Bowman S. L., Atanasov N., Daniilidis K., Pappas G. J.** Probabilistic data association for semantic slam // Proc. of the IEEE Intern. Conf. on Robotics and Automation (ICRA 2017). Singapore, Singapore, 29 May – 3 June, 2017. P. 1722–1729.
7. **Civera J., Galvez-Lopez D., Riazuelo L. et al.** Towards semantic SLAM using a monocular camera // Proc. of the IEEE/RSJ Intern. Conf. on Intelligent Robots and Systems (IROS 2011). San Francisco, USA, 25–30 Sept., 2011. P. 1277–1284.
8. **Klein G., Murray D.** Parallel tracking and mapping for small AR workspaces // Proc. of the 6th IEEE and ACM Intern. Symposium on Mixed and Augmented Reality (ISMAR 2007). Nara, Japan. 13–16 Nov., 2007. P. 225–234.
9. **Newcombe R. A., Lovegrove S. J., Davison A. J.** Dense tracking and mapping in real-time // Proc. of the IEEE Intern. Conf. on Computer Vision (ICCV 2011). Barcelona, Spain, 6–13 Nov., 2011. P. 2320–2327.
10. **Кирильченко А. А., Платонов А. К., Соколов С. М.** Теоретические аспекты организации интерпретирующей навигации мобильного робота. М., 2002. 40 с. (Препр. /Ин-т прикл. математики им. М. В. Келдыша РАН; № 5).
11. **Chaumette F., Hutchinson S.** Visual servo control. I. Basic approaches // IEEE Robot. Automat. Mag. 2006. **13**, N 4. P. 82–90.
12. **Jia B., Liu S.** Switched visual servo control of nonholonomic mobile robots with field-of-view constraints based on homography // Control Theory Technol. 2015. **13**, N 4. P. 311–320.
13. **Юдин Д. А., Проценко В. В., Постолюнский Г. Г. и др.** Система технического зрения для автоматического ориентирования и позиционирования мобильного робота // Робототехника и техническая кибернетика. 2014. **2**, № 1. С. 70–75.
14. **Барамия Д. А., Дьяков М. С., Кузиковский С. А., Лаврентьев М. М.** Система одновременной локализации и построения карты на основе подхода CoreSLAM // Автометрия. 2017. **53**, № 6. С. 77–82.
15. **Яковлев К. С., Хитыков В. В., Логинов М. И., Петров А. В.** Система навигации группы БЛА на основе маркёров // Робототехника и техническая кибернетика. 2014. **5**, № 4. С. 44–48.
16. **Bay H., Tuytelaars T., Van Gool L.** Surf: Speeded up robust features // Proc. of the Eur. Conf. on Computer Vision (ICCV 2006). Graz, Austria, 7–13 May, 2006. P. 404–417.
17. **Lowe D. G.** Object recognition from local scale-invariant features // Proc. of the 7th IEEE Intern. Conf. on Computer Vision (ICCV 1999). Kerkyra, Greece, 20–27 Sept., 1999. Vol. 2. P. 1150–1157.
18. **Calonder M., Lepetit V., Strecha C. et al.** Brief: Binary robust independent elementary features // Proc. of the Eur. Conf. on Computer Vision (ICCV 2010). Crete, Greece, 5–11 Sept., 2010. P. 778–792.
19. **Rublee E., Rabaud V., Konolige K., Bradski G.** ORB: An efficient alternative to SIFT or SURF // Proc. of the IEEE Intern. Conf. on Computer Vision (ICCV 2011). Barcelona, Spain, 6–13 Nov., 2011. P. 2564–2571.
20. **Fischler M. A., Bolles R. C.** Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography // Commun. ACM. 1981. **24**, N 6. P. 381–395.
21. **Nistér D., Naroditsky O., Bergen J.** Visual odometry for ground vehicle applications // Journ. Field Robotics. 2006. **23**, N 1. P. 3–20.

22. **Макаров И. М.** Принципы построения интеллектуальных регуляторов и систем управления средствами ВВТ // Изв. ЮФУ. Технические науки. 2006. **58**, № 3. С. 3–7.
23. **Tolman E. C.** Cognitive maps in rats and men // Psych. Rev. 1948. **55**, N 4. P. 189.
24. **Vasudevan S., Gächter S., Nguyen V., Siegwart R.** Cognitive maps for mobile robots an object based approach // Robotics and Autonomous Syst. 2007. **55**, N 5. P. 359–371.
25. **Garcia-Garcia A., Orts-Escolano S., Oprea S. et al.** A review on deep learning techniques applied to semantic segmentation // Comp. Vis. Pattern Recogn. Cornell Univers., 2017. URL: <https://arXiv preprint arXiv:1704.06857> (дата обращения: 10.03.2018).
26. **Krizhevsky A., Sutskever I., Hinton G. E.** Imagenet classification with deep convolutional neural networks // Proc. of the 26th Conf. on Neural Information Processing Systems. Lake Tahoe, USA, 3–8 Dec., 2012. P. 1097–1105.
27. **Simonyan K., Zisserman A.** Very deep convolutional networks for large-scale image recognition // Comp. Vis. Pattern Recogn. Cornell Univers., 2015. URL: <https://arXiv preprint arXiv:1409.1556> (дата обращения: 02.05.2018).
28. **Szegedy C., Liu W., Jia Y., Sermanet P.** Going deeper with convolutions // Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2015). Boston, USA, 7–12 June, 2015. P. 1–9.
29. **He K., Zhang X., Ren S., Sun J.** Deep residual learning for image recognition // Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2015). Boston, USA, 7–12 June, 2016. P. 770–778.
30. **Ren S., He K., Girshick R., Sun J.** Faster R-CNN: Towards real-time object detection with region proposal networks // Proc. of the 29th Conf. on Neural Information Processing Systems. Montreal, Canada, 7–12 Dec., 2015. P. 91–99.
31. **Liu W., Anguelov D., Erhan D., Szegedy C.** Ssd: Single shot multibox detector // Proc. of the Eur. Conf. on Computer Vision (ICCV 2016). Amsterdam, The Netherlands, 8–16 Oct., 2016. P. 21–37.
32. **Redmon J., Divvala S., Girshick R., Farhadi A.** You only look once: Unified, real-time object detection // Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2016). Las Vegas, USA, 27–30 June, 2016. P. 779–788.
33. **Long J., Shelhamer E., Darrell T.** Fully convolutional networks for semantic segmentation // Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2015). Boston, USA, 7–12 June, 2015. P. 3431–3440.
34. **Ronneberger O., Fischer P., Brox T.** U-net: Convolutional networks for biomedical image segmentation // Proc. of the Intern. Conf. on Medical Image Computing and Computer-Assisted Intervention. Munich, Germany, 5–9 Oct., 2015. P. 234–241.
35. **Badrinarayanan V., Kendall A., Cipolla R.** Segnet: A deep convolutional encoder-decoder architecture for image segmentation // IEEE Trans. Pattern Analysis Machine Intell. 2017. **39**, N 12. P. 2481–2495.
36. **Pinheiro P. O., Collobert R., Dollár P.** Learning to segment object candidates // Proc. of the 29th Conf. on Neural Information Processing Systems. Montreal, Canada, 7–12 Dec., 2015. P. 1990–1998.
37. **Ющенко А. С.** Маршрутизация движения мобильного робота в условиях неопределённости // Мехатроника, автоматизация, управление. 2004. № 1. С. 31–38.
38. **Thrun S., Burgard W., Fox D.** A probabilistic approach to concurrent mapping and localization for mobile robots // Autonomous Robots. 1998. **5**, N 3–4. P. 253–271.
39. **Song S., Yu F., Zeng A., Chang A. X.** Semantic scene completion from a single depth image // Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2017). Honolulu, USA, 21–26 July, 2017. P. 190–198.

40. **Dai A., Nießner M.** 3DMV: Joint 3D-Multi-View Prediction for 3D Semantic Scene Segmentation // *Comp. Vis. Pattern Recogn. Cornell Univers.*, 2018. URL: <https://arXiv preprint arXiv:1803.10409> (дата обращения: 25.11.2018).
41. **Cherabier I., Schonberger J. L., Oswald M. R. et al.** Learning priors for semantic 3D reconstruction // *Proc. of the Eur. Conf. on Computer Vision (ECCV 2018)*. Munich, Germany, 8–14 Sept., 2018. Vol. 2. P. 314–330.
42. **Dai A., Chang A. X., Savva M., Halber M.** ScanNet: Richly-Annotated 3D Reconstructions of Indoor Scenes // *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2017)*. Honolulu, USA, 21–26 July, 2017. Vol. 2. P. 2432–2443.
43. **Lawin F. J., Danelljan M., Tosteberg P.** Deep projective 3D semantic segmentation // *Proc. of the Intern. Conf. on Computer Analysis of Images and Patterns*. Ystad, Sweden, 22–24 Aug., 2017. P. 95–107.
44. **McCormac J., Handa A., Davison A., Leutenegger S.** Semanticfusion: Dense 3d semantic mapping with convolutional neural networks // *Proc. of the IEEE Intern. Conf. on Robotics and Automation (ICRA 2017)*. Singapore, Singapore, 29 May – 3 June, 2017. P. 4628–4635.
45. **Qi C. R., Su H., Mo K., Guibas L. J.** Pointnet: Deep learning on point sets for 3d classification and segmentation // *Proc. of the IEEE Conf. Computer Vision and Pattern Recognition (CVPR 2017)*. Honolulu, USA, 21–26 July, 2017. P. 77–85.
46. **Hua B. S., Tran M. K., Yeung S. K.** Pointwise convolutional neural networks // *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2018)*. Salt Lake City, USA, 18–23 June, 2018. P. 984–993.
47. **Qi C. R., Yi L., Su H., Guibas L. J.** Pointnet++: Deep hierarchical feature learning on point sets in a metric space // *Proc. of the 29th Conf. on Neural Information Processing Systems*. Long Beach, USA, 4–9 Dec., 2017. P. 5099–5108.
48. **Pham Q. H., Hua B. S., Nguyen D. T., Yeung S. K.** Real-time Progressive 3D Semantic Segmentation for Indoor Scene // *Comp. Vis. Pattern Recogn. Cornell Univers.*, 2017. URL: <https://arXiv preprint arXiv:1804.00257> (дата обращения: 25.01.2019).

Поступила в редакцию 21.11.2018

После доработки 23.01.2019

Принята к публикации 24.01.2019
