

УДК 004.896

doi: 10.17586/2226-1494-2019-19-4-622-629

ЛОКАЛИЗАЦИЯ МОБИЛЬНОГО РОБОТА С ФИЛЬТРОМ ЧАСТИЦ ПРИ ОБНАРУЖЕНИИ И СЕГМЕНТАЦИЕЙ ОБЪЕКТОВ

М.И. Евстигнеев^a, Ю.В. Литвинов^a, В.В. Мазулина^{a,b}

^a Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

^b Технический университет Эйндховена, Эйндховен, 5600MB, Нидерланды

Адрес для переписки: mevstigneev@corp.ifmo.ru

Информация о статье

Поступила в редакцию 07.05.19, принята к печати 30.05.19

Язык статьи — русский

Ссылка для цитирования: Евстигнеев М.И., Литвинов Ю.В., Мазулина В.В. Локализация мобильного робота с фильтром частиц при обнаружении и сегментацией объектов // Научно-технический вестник информационных технологий, механики и оптики. 2019. Т. 19. № 4. С. 622–629. doi: 10.17586/2226-1494-2019-19-4-622-629

Аннотация

Предмет исследования. Предложен метод локализации мобильного робота с использованием фильтра частиц (метода Монте-Карло), основанный на компьютерном зрении. Алгоритм локализации использует отличительные ориентиры, которые понятны человеку. Семантическая информация используется в модели движения с данными о дальности и без них. **Метод.** Принцип работы модифицированного алгоритма локализации заключается в использовании семантических подсказок высокого уровня. Вместо выдавливания плана этажа в третье измерение производится свертка трехмерного мира в двумерное представление и выборка дискриминационных ориентиров высокого уровня. Этот подход используется для представления глобальной локализации, которая опирается исключительно на семантические метки, присутствующие в плане этажа и извлеченные из изображений RGB. **Основные результаты.** В работе продемонстрировано, что локализация с сегментацией объектов, основанная на отличительных ориентирах, является эффективной альтернативой традиционному сканированию. Исследование производится в наборе данных плана этажа, а также проводится сравнение нескольких подходов с точки зрения качественной и количественной оценки локализации на уровне комнаты и глобальной локализации. Продемонстрировано, что семантическая информация дополняет современные методы, обеспечивая снижение ошибок до 35 %. **Практическая значимость.** Представлена новая структура восприятия и локализации, которая использует семантические данные и информацию о расстояниях. Новая платформа может быть использована для локализации как превосходящая традиционные алгоритмы, основанные на методе Монте-Карло.

Ключевые слова

локализация робота, SLAM, техническое зрение, фильтр частиц, нейронные сети

doi: 10.17586/2226-1494-2019-19-4-622-629

LOCALIZATION OF MOBILE ROBOT WITH PARTICLE FILTER AT DETECTION AND SEGMENTATION OF OBJECTS

M.I. Evstigneev^a, Yu.V. Litvinov^a, V.V. Mazulina^{a,b}

^a ITMO University, Saint Petersburg, 197101, Russian Federation

^b Technical University of Eindhoven, Eindhoven, 5600MB, Netherlands

Corresponding author: mevstigneev@corp.ifmo.ru

Article info

Received 07.05.19, accepted 30.05.19

Article in Russian

For citation: Evstigneev M.I., Litvinov Yu.V., Mazulina V.V. Localization of mobile robot with particle filter at detection and segmentation of objects. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2019, vol. 19, no. 4, pp. 622–629 (in Russian). doi: 10.17586/2226-1494-2019-19-4-622-629

Abstract

Subject of Research. The paper presents a method for mobile robot localization using a particle filter (Monte-Carlo method) based on computer vision. The localization algorithm uses distinctive landmarks that are understandable to a man. Semantic information is used in the motion model with and without range data. **Method.** The operation principle of the modified localization algorithm lies in applying high-level semantic prompts. Instead of squeezing the floor plan into the third dimension, the three-dimensional world is convolved into a two-dimensional representation and a sample of high-level discriminatory

landmarks. This approach is used to represent global localization, which relies exclusively on semantic labels present in the floor plan and extracted from RGB images. **Main Results.** We demonstrate that localization with segmentation of objects, based on distinctive landmarks, is an effective alternative to traditional scanning. The study is performed in a floor plan data set, and several approaches are compared in terms of qualitative and quantitative localization at room level and global localization. It is shown that semantic information complements modern methods, ensuring that errors are reduced to 35 %. **Practical Relevance.** We have presented a new structure of perception and localization which uses semantic data and information about distances. The new platform can be used for localization as superior to traditional algorithms based on the Monte Carlo method.

Keywords

robot localization, SLAM, technical vision, particle filter, neural networks

Введение

В настоящее время современные роботы стали повсеместным явлением и выполняют сложные или утомительные задачи, будь то складская сортировка, уборка, помощь в поисково-спасательных операциях, видеосъемка и 3D-реконструкция. Локализация в здании является важным аспектом их работы и позволяет взаимодействовать с миром. Традиционные системы SLAM (simultaneous localization and mapping) могут обеспечить локализацию на карте, которая создается на «лету», но они подвержены дрейфу с точки зрения положения и масштаба и некорректным замыканием циклов движения, поэтому используются системы локализации для связи позиции робота с уже существующей картой. В основанных на техническом зрении подходах локализации обычно используются эвристические алгоритмы для преобразования 2D-плана в 3D-систему координат визуального SLAM. Примеры включают работы [1], в которых используются визуальные подсказки. Общая проблема с этими подходами состоит в том, что трехмерные данные, извлеченные из изображения, обычно ортогональны плану этажа. Это означает, что должны быть сделаны предположения относительно размеров, отсутствующих в плане, и эти подходы также игнорируют семантическую информацию.

Локализация с фильтром частиц, называемая также локализацией Монте-Карло (ЛМК) [2], улучшила методы на основе фильтра Калмана, позволив представить мультимодальные распределения. Эти подходы являются надежными и до сих пор считаются самыми современными во многих роботизированных приложениях, однако для надежной работы эти подходы требуют дорогостоящих датчиков, таких как лидар. Вместо этого в [3] расширили подход к использованию сенсорных моделей на основе технического зрения. С ростом популярности датчиков RGB-D стали возможны более надежные подходы ЛМК, основанные на техническом зрении, так в [4] использовали визуальную одометрию и уже существующие карты. Исследования [5] ближе всего похожи на разрабатываемый метод. В них производятся частичные реконструкции экструдированного плана этажа, и модель ЛМК основана на сопоставлениях с этими реконструкциями.

Рассматриваемый подход схож с методом в работах [6, 7], где угловое искажение известных ориентиров может использоваться для определения местоположения роботизированного агента. Однако рассматриваемый метод отличается от них, поскольку он не требует активных ориентиров с известными позициями. Вместо этого предлагается использовать семантическую информацию, которая уже существует в мире, и использовать угловое распределение обнаруженных семантических меток для локализации робота.

Развитие методов семантической сегментации [8, 9] привело к возможности ее использования при сопоставлении изображений с картами помещений в задаче локализации. В работе [10] рассматривается метод, который использует основанные на сверточных нейронных сетях (CNN) предсказания глубины и семантические метки [11], применяющийся для локализации. В текущей работе предлагается использовать основанную на CNN семантическую сегментацию для извлечения меток, которые по своей природе присутствуют в этажах.

Семантическая планировка этажей

Для ЛМК требуется план этажа и/или ранее созданная карта расстояний, которая точна в масштабе и согласована на глобальном уровне. Предлагается использование читаемых человеком планов этажей, что делает систему более широко применимой. Различия между планом этажа и наблюдениями робота (неточности, масштаб и мебель), очевидно, негативно сказываются на точности методов локализации. Чтобы преодолеть это, мы дополняем локализацию семантическими метками, извлеченными из плана этажа. Мы ограничиваемся стенами, дверями и окнами (рис. 1, а), которые легко извлекаются из плана этажа, а также характерны для локализации человека. Чтобы сделать помеченный план этажа читаемым для робота, он должен быть сначала преобразован в сетку занятости (рис. 1, б). Сетка занятости — двумерное представление мира, в котором каждая ячейка имеет вероятность занятости, определяемую ее нормализованным значением шкалы серого.

Если M представляет собой набор 2D-позиций, карту можно определить как $V = \{v_m; m \in M \subset Z^{+2}\}$. Тогда, предполагая, что $L = \{a, d, w\}$ — это набор возможных меток ячеек (стена, дверь, окно), каждая ячейка определяется как $v_m = \langle v_m^0, v_m^w, v_m^d, v_m^a \rangle$, где v_m^0 — это вероятность занятости, $l \in L$ обозначает вероятность метки.

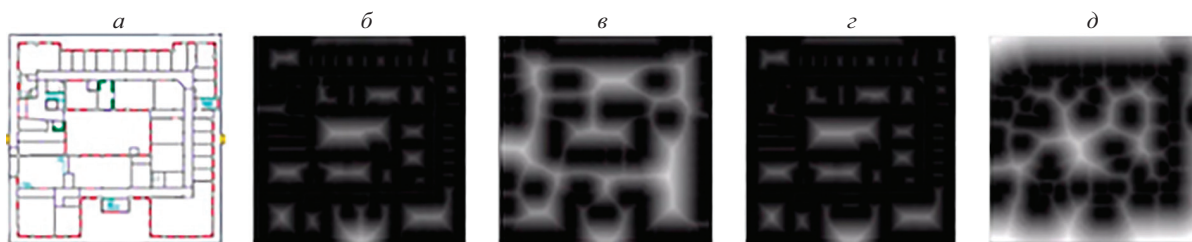


Рис. 1. План этажа и сетки занятости: семантический план этажа (а), сетка занятости (б), сетка занятости окна (в), сетка занятости стены (г), сетка занятости двери (д)

Робототехнические системы с RGB-D камерами работают с массивом расстояний (r_t^k) и углов (θ_t^k). В данном методе мы добавляем семантическую метку (l_t^k) к этому массиву. Вместо одновременного использования всего изображения массивы располагаются вдоль горизонтальных линий сканирования ($z_t = \{\langle \theta_t^k, r_t^k, l_t^k \rangle; k = 1..K\}$), где k — это горизонтальное положение пикселя. В работе предполагается, что центральная линия развертки параллельна плоскости земли и, следовательно, используется для свертывания трехмерной информации изображения RGB-D в двухмерный план этажа.

Семантическое обнаружение — структура, которая сочетает возможности семантического восприятия с ЛМК. На рис. 2 показан пример сканирования в плане этажа. Можно использовать любой подход семантической сегментации, однако подходы, основанные на глубоком обучении, в настоящее время доминируют по эффективным показателям. Поэтому используется сеть кодера-декодера на основе CNN [12]. Сеть обучается на наборе данных [13] и может определять двери, стены, полы, потолки, мебель и окна. Метка l_t^k — это метка в пикселе k вдоль горизонтальной линии сканирования.

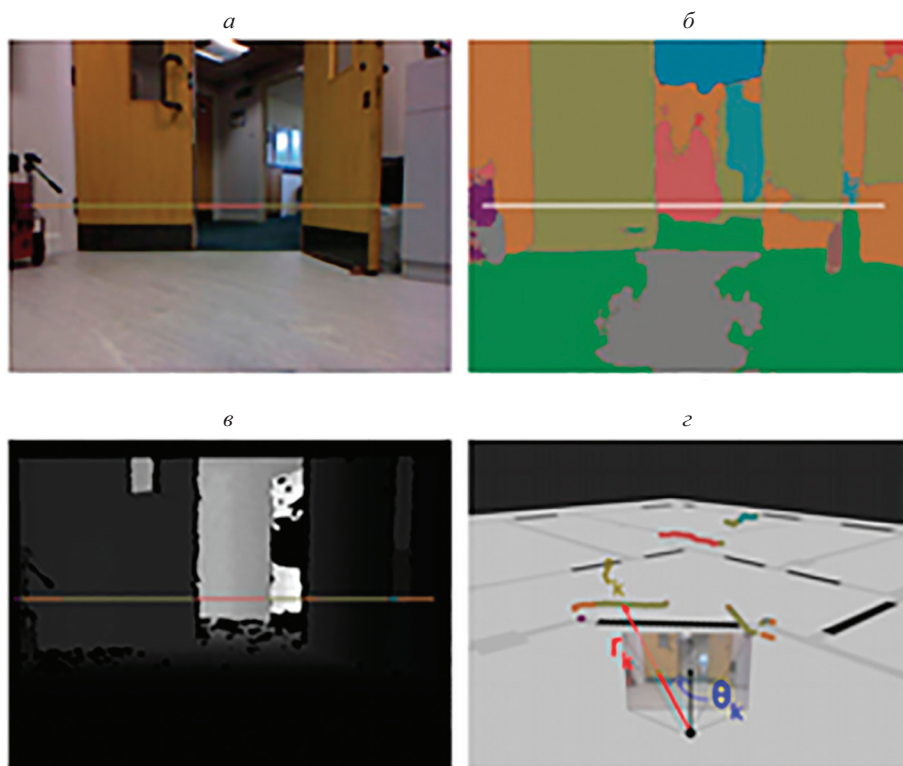


Рис. 2. Семантическая маркировка изображений: изображение RGB (а), семантическая маркировка CNN (б), карта глубины (в), пример сканирования в плане этажа (г)

Модели движения ЛМК обычно представлены распределением $\Pr(s_t^i | u_t, s_{t-1}^i)$, где предыдущий набор частиц s_{t-1}^i распространяется с использованием измерений одометрии u_t в текущий набор частиц s_t^i . Однако также хорошо изучено фактическое распределение, которое аппроксимируется как $\Pr(s_t^i | u_t, s_{t-1}^i, V)$. Это кодирует идею о том, что определенные движения более или менее вероятны в зависимости от карты (например, движения сквозь стены). Предполагая, что движение робота мало, это можно выразить формулой:

$$\Pr(s_t^i | u_t, s_{t-1}^i, V) = \kappa \Pr(s_t^i | u_t, s_{t-1}^i) \Pr(s_{t-1}^i | V),$$

(например, в работе [14]), где κ — нормализующий фактор, а V — множество, содержащее каждую ячейку на карте. Это позволяет рассматривать две вероятности независимо друг от друга. Движение $\Pr(s_i^r | u_t, s_{t-1}^i)$ определяется так же, как в [15]. Приоритет — это вероятность заполнения ячейки, которая содержит s_i^i , что есть $\Pr(s_{t-1}^i | V) = 1 - \Pr(v_{s_{t-1}}^0)$.

Однако этот подход предварительной оценки становится проблематичным при использовании созданных человеком планов этажей, поскольку в них обычно присутствуют артефакты изображения, возникающие в процессе сканирования. Поэтому порог занятости выражается:

$$\Pr(v_{s_{t-1}}^0) = \begin{cases} 1 & \text{если } v_{s_{t-1}}^0 \geq \tau_0 \\ 0 & \text{в противном случае} \end{cases}, \quad (1)$$

где τ_0 — определенный пользователем порог. Это усугубляет проблемы с точностью и перекрытием плана этажа и создает реальные проблемы, когда частицы движутся через двери, так как многие действительные частицы будут выброшены при контакте с ожидаемым краем двери. Вместо этого мы предлагаем дополнить это побочным фактором (ϵ_G), который позволяет частицам иметь больше свободы. Поэтому предлагается:

$$\Pr(s_{t-1}^i | V) = (1 - \Pr(v_{s_{t-1}}^0)) e^{-\epsilon_G \delta_a},$$

где δ_a — расстояние до ближайшей двери.

Что еще более важно, ϵ_G является пользовательским фактором, который определяет, насколько строго применяется это наказание. Настройка $\epsilon_G = 0$ позволяет частицам перемещаться через стены без каких-либо штрафов, в то время как очень высокие значения приближают к уравнению (1). Эта модель более вероятно точна, чем модель занятости, используемая в большинстве подходов ЛМК, и имеет дополнительное преимущество использования семантической информации.

Модель поля правдоподобия рассчитывает карту расстояний. Для каждой ячейки v_m расстояние до ближайшей занятой ячейки можно выразить как:

$$\delta_0(m) = \min_{m'} \|m - m'\|, \quad v_m^0 > \tau_0.$$

Когда принимается измерение $z_t^k = \langle \theta_t^k, r_t^k \rangle$, конечная точка оценивается и используется в качестве индекса для карты расстояний. Предполагая распределение ошибок по Гауссу, вес каждой частицы s_t^i может быть оценен как:

$$\Pr_{\text{Range}}(z_t^k | s_t^i, V) = e^{-\delta_0^2 / 2\sigma_0^2}, \quad (2)$$

где δ_0 — это значение, полученное из карты расстояний, а σ_0 определяется шумовыми характеристиками датчика. Эта модель имеет ограничения, так как не использует семантическую информацию, параметр σ_0 должен быть оценен пользователем. Предполагается, что все измерения в пределах сканирования имеют одинаковые параметры шума, и параметр σ_0 не способен работать при отсутствии измерений дальности.

Для каждой метки, присутствующей в плане этажа, мы можем рассчитать карту расстояний, в которой хранится кратчайшее расстояние до ячейки с такой же меткой. Для каждой ячейки карты v_m мы можем оценить расстояние до ближайшей ячейки каждой метки как:

$$\delta_l(m) = \min_{m'} \|m - m'\|, \quad v_m^l > \tau_0.$$

где $\delta_l = \{\delta_a, \delta_d, \delta_w\}$ — расстояния до ближайшей стены, двери и окна соответственно. На рис. 1 показаны карты расстояний для каждой метки.

Когда мы получаем наблюдение $z_t^k = \langle \theta_t^k, r_t^k, l_t^k \rangle$, мы используем информацию об углах θ_t^k и расстоянии r_t^k для оценки конечной точки сканирования. Затем мы используем метку l_t^k , чтобы решить, какое поле семантической вероятности использовать. Используя конечную точку из предыдущего шага, вероятность метки можно оценить аналогично уравнению (2):

$$\Pr_{\text{Label}}(z_t^k | s_t^i, V) = e^{-\delta_l^2 / 2\sigma_l^2}, \quad (3)$$

где δ_l — расстояние до ближайшей ячейки соответствующей метки, σ_l — стандартное отклонение, которое мы определим с помощью предыдущей метки. Вероятность наблюдения с учетом карты и позиции может быть оценена:

$$\Pr(z_t^k | s_t^i, V) = \epsilon_0 \Pr_{\text{Range}}(z_t^k | s_t^i, V) + \epsilon_l \Pr_{\text{Label}}(z_t^k | s_t^i, V),$$

где ϵ_0 и ϵ_l — определенные пользователем веса. Когда $\epsilon_l = 0$, вероятность такая же, как у стандартной ЛМК. С другой стороны, когда $\epsilon_0 = 0$, подход использует только семантическую информацию. В отличие

от сканеров дальности σ_l не может быть связан с физическими свойствами датчика. Данное стандартное отклонение оценивается непосредственно от предыдущей метки на карте. Определение σ_l этим способом имеет преимущество в том, что не требует настройки.

Когда человек читает план этажа, уникальные ориентиры являются наиболее отличительными чертами: легче локализоваться на плане этажа из конфигурации дверей и окон, чем из конфигурации стен. Это приводит к простому пониманию: более низкие приоритеты являются более дискриминационными. Следовательно, σ_l привязан к каждой предыдущей метке из-за того, что он настраивается на один параметр меньше и неявно делает наблюдение редких ориентиров более выгодным.

Отношение σ_l к метке, предшествующей $\Pr(I)$, контролирует насколько плавно распределение распадается с учетом расстояния от клетки. Чем меньше $\Pr(I)$, тем плавнее затухание. По сути алгоритм локализации должен быть более снисходительным к редким меткам.

Преимуществом подхода является способность выполнять описанную методологию при полном отсутствии измерений дальности. До сих пор мы формализовали этот подход в предположении, что были получены два вида массивов: либо массивы $\langle \theta_i^k, r_i^k \rangle$ (существующие подходы), либо массивы $\langle \theta_i^k, r_i^k, l_i^k \rangle$ (подход на основе семантической локализации). Тем не менее, этот подход способен работать напрямую с массивами $\langle \theta_i^k, l_i^k \rangle$.

Работа без измерения расстояний проста. Модели рейкастинга и правдоподобия объединяются в новом подходе, который позволяет избежать вырождений, которые могут возникнуть в традиционных подходах ЛМК. В стандартном подходе операция рейкастинга завершается, когда достигается занятая ячейка и оценивается вероятность

$$\Pr(z_i^k | s_i', V) = e^{-\frac{(r_i^k - r_i^{k*})^2}{2\sigma_0^2}},$$

где r_i^k — это расстояние, полученное от датчика, r_i^{k*} — расстояние, пройденное лучом. К сожалению, в отсутствие измерения на основе диапазона r_i^k это невозможно. Использование стандартной карты расстояний также невозможно, так как мы не можем оценить конечную точку луча. Использование рейкастинга в карте расстояний не выполняется аналогичным образом. Рейкастинг заканчивается на занятой ячейке, подразумевая $\delta_0 = 0$ для каждого луча.

С другой стороны, все еще могут использоваться семантические поля занятости, так как σ_l будет по-прежнему иметь значимое и различающее значение. Тогда выполняется рейкастинг для каждого. Однако вместо сравнения r_i^k и r_i^{k*} или использования δ_0 , метка l_i^k определяет, какое поле вероятности использовать. Тогда функция стоимости:

$$\Pr(z_i^k | s_i', V) = \Pr_{Label}(z_i^k | s_i', V), \quad (4)$$

где $\Pr_{Label}(z_i^k | s_i', V)$ определяется в соответствии с уравнением (3). Этот метод представляет собой комбинацию лучевой модели и модели занятости. При отсутствии измерений дальности для оценки конечной точки, этот гибридный подход использует семантический рейкастинг для нахождения ближайшей занятой ячейки. Затем расстояния используются для обеспечения гладкости уравнения (4), из чего следует, что вероятность наблюдения прямо пропорциональна угловому распределению меток.

Результаты

Траектория робота находится в той же плоскости, что и план этажа, и фиксируется с помощью камеры RGB-D, чтобы извлечь все возможные комбинации массивов (расстояние, угол и метка). Используется набор данных плана этажа, изображенный на рис. 1, а. Набор данных был собран с использованием платформы робота с сенсором Kinect.

Обычно реальная траектория для локализации на плане этажа оценивается вручную [16] или с использованием систем MotionCapture [17]. Однако ручная оценка является трудоемкой и непрактичной, а MotionCapture – дорогостоящая и трудно калибруемая система. Чтобы преодолеть эти ограничения, используется хорошо зарекомендовавший себя алгоритм RGB-D SLAM [15], который обеспечивает очень точную оценку позиционирования.

Для количественной оценки представленного подхода по отношению к реальным значениям используется метрика ошибки траектории, представленная в работе [7]. Эта оценка выполняется путем регистрации двух траекторий [16]. Среднеквадратическая ошибка (СКО) метрики указывает на производительность (таблица).

Мы сравниваем алгоритм с популярным подходом ЛМК, представленным в операционной системе Robot Operating System (ROS), называемым адаптивной локализацией Монте-Карло (AMCL) [2]. Хотя существуют более современные подходы [17], они основаны на тех же принципах, что и AMCL, и просто меняют стратегию отбора частиц. Во всех экспериментах параметры (такие как) сохраняются одинаковыми. Изменяются только параметры $\epsilon_b, \epsilon_0, \epsilon_G$.

Таблица. Средняя ошибка траектории

Алгоритм	СКО, м	Среднее, м	Мин., м	Макс., м
AMCL	0,25	0,22	0,05	0,96
Метки и глубина	0,20	0,15	0,03	0,56
Рейкастинг (метки) ($\epsilon_G = 3$)	0,41	0,28	0,08	1,52
Рейкастинг (метки) ($\epsilon_G = 7$)	0,59	0,39	0,03	2,24

Для этой оценки AMCL и предлагаемому подходу дается инициализация на уровне комнаты со стандартными отклонениями 2,0 м в (x, y) и 2,0 рад в θ . Системы работали с диапазоном количества частиц от 250 до 1000. Ошибка записывается при добавлении каждого нового изображения в набор данных. С точки зрения качественной оценки показывается поведение сходимости и оценочный путь. Поведение сходимости можно увидеть на рис. 3. На рис. 3, *a* показано распределение частиц в комнате, в которой находится робот. Когда робот начинает двигаться, мы можем видеть, как сходятся AMCL (рис. 3, *б*), версия, основанная на расстоянии семантической локализации (рис. 3, *в*), и лучевая версия (рис. 3, *г*).

Несмотря на то что лучевой подход имеет предсказуемо большую дисперсию частиц, фильтр успешно локализован, так как восстановленное облако точек Kinect правильно выровнено с планом этажа. Важно отметить, что, хотя облако точек Kinect присутствует для визуализации в лучевом методе, оно не используется.

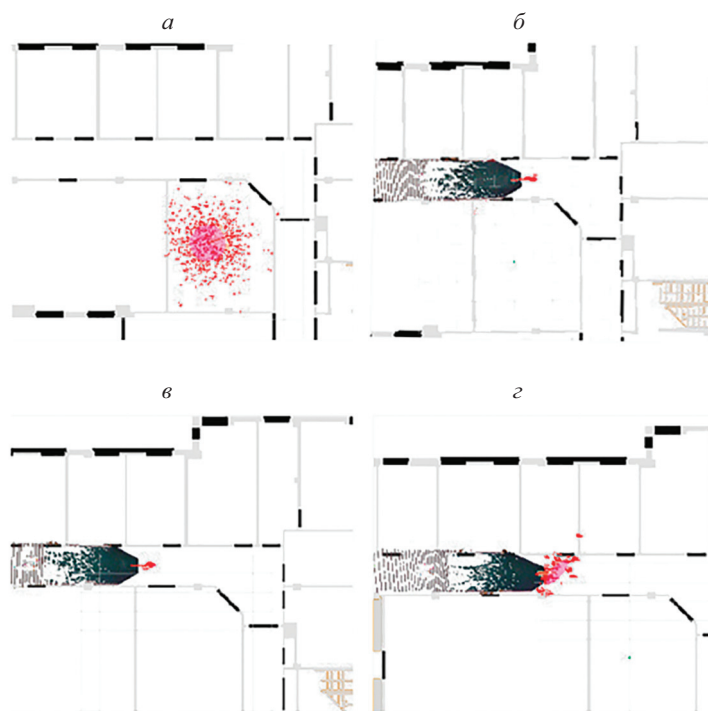


Рис. 3. Качественная оценка локализации: инициализация на уровне комнаты (*a*), AMCL (*б*), глубина и метки (*в*), только метки (*г*)



Рис. 4. Расчетный путь от инициализации на уровне комнаты: AMCL (*a*), глубина и метки (*б*), только метки (*в*)

Расчетные пути можно увидеть на рис. 4, где красный путь — это оценочный путь, а зеленый — истинный путь. На рис. 4, а показан метод AMCL, который пытается сойтись в начале последовательности. На рис. 4, б изображено, что подход, основанный на объединении меток и расстояний, сходится быстрее и поддерживает производительность, аналогичную AMCL. Он лишь слегка отклоняется от пути в конце двусмысленного коридора слева, что также происходит в AMCL. Лучевой подход работает стабильно, хотя для сходимости требуется больше времени, что видно по оценочной траектории на рис. 4, в. Он корректируется и отклоняется от пути только в областях большой неопределенности (например, в длинных коридорах).

Заключение

В представленной работе введена идея семантического плана этажа с характерными и понятными для людей ориентирами. Представляется новый способ восприятия, который добавляет семантические метки к традиционной информации, полученной от измерителей дальности. Затем эти идеи используются в новом подходе локализации, основанном на локализации Монте-Карло. Этот подход позволяет использовать семантическую информацию, представленную на карте, для определения новой модели движения. Он также способен использовать метки из сегментации на основе CNN для локализации на карте. Подход работоспособен как при наличии, так и при отсутствии измерений дальности. Эксперименты показывают, что новая семантическая информация в значительной степени дополняет современные методы, обеспечивая снижение ошибок до 35 %. Основываясь на этом, можно заключить, что применение семантической информации должно быть дополнительно изучено в более широкой области робототехники.

Литература

1. Liu C., Schwing A.G., Kundu K., Urtasun R., Fidler S. Rent3D: Floor-plan priors for monocular layout estimation // Proc. IEEE Conf. on Computer Vision and Pattern Recognition. 2015. doi: 10.1109/CVPR.2015.7298963
2. Thrun S., Fox D., Burgard W., Dellaert F. Robust Monte Carlo localization for mobile robots // Artificial Intelligence. 2001. V. 128. N 1-2. P. 99–141. doi: 10.1016/S0004-3702(01)00069-8
3. Dellaert F. Using the condensation algorithm for robust, vision-based mobile robot localization // Proc. IEEE Conf. on Computer Vision and Pattern Recognition. 1999. P. 10–12. doi: 10.1109/CVPR.1999.784976
4. Brubaker M.A., Geiger A., Urtasun R. Lost! Leveraging the crowd for probabilistic visual self-localization // Proc. IEEE Conf. on Computer Vision and Pattern Recognition. 2013. P. 10–16. doi: 10.1109/CVPR.2013.393
5. Chu H., Kim D.K., Chen T. You are here: mimicking the human thinking process in reading floor-plans // Proc. IEEE Int. Conf. on Computer Vision. 2015. doi: 10.1109/ICCV.2015.255
6. Briechele K., Hanebeck U.D. Localization of a mobile robot using relative bearing measurements // IEEE Transactions on Robotics and Automation. 2002. V. 20. N 1. P. 36–44. doi: 10.1109/TRA.2003.820933
7. Thrun S. Probabilistic robotics // Communications of the ACM. 2002. V. 45. N 3. doi: 10.1145/504729.504754
8. Badrinarayanan V. SegNet: a deep convolutional encoder-decoder architecture for image segmentation // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2017. V. 39. N 13. P. 2481–2495. doi: 10.1109/TPAMI.2016.2644615
9. Shelhamer E., Long J., Darrell T. Fully convolutional networks for semantic segmentation // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2017. V. 39. N 13. P. 640–651. doi: 10.1109/TPAMI.2016.2572683
10. Tateno K., Tombari F., Laina I., Navab N. CNN-SLAM: Real-time dense monocular SLAM with learned depth prediction // Proc. IEEE Conf. on Computer Vision and Pattern Recognition. 2017. doi: 10.1109/cvpr.2017.695
11. Laina I., Ruppel C., Belagiannis V. Deeper depth prediction with fully convolutional residual networks // Proc. 4th Int. Conf. on 3D Vision. 2016. doi: 10.1109/3dv.2016.32
12. Kendall A., Badrinarayanan V., Cipolla R. Bayesian SegNet: model uncertainty in deep convolutional encoder-decoder architectures for scene understanding // Proceedings of the British Machine Vision Conference. 2017. doi: 10.5244/c.31.57
13. Xiao J., Owens A., Torralba A. SUN3D: A database of big spaces reconstructed using SfM and object labels // Proc. IEEE Int. Conf. on Computer Vision. 2013. doi: 10.1109/ICCV.2013.458

References

1. Liu C., Schwing A.G., Kundu K., Urtasun R., Fidler S. Rent3D: Floor-plan priors for monocular layout estimation. *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 2015. doi: 10.1109/CVPR.2015.7298963
2. Thrun S., Fox D., Burgard W., Dellaert F. Robust Monte Carlo localization for mobile robots. *Artificial Intelligence*, 2001, vol. 128, no. 1-2, pp. 99–141. doi: 10.1016/S0004-3702(01)00069-8
3. Dellaert F. Using the condensation algorithm for robust, vision-based mobile robot localization. *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 1999, pp. 10–12. doi: 10.1109/CVPR.1999.784976
4. Brubaker M.A., Geiger A., Urtasun R. Lost! Leveraging the crowd for probabilistic visual self-localization. *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 2013, pp. 10–16. doi: 10.1109/CVPR.2013.393
5. Chu H., Kim D.K., Chen T. You are here: mimicking the human thinking process in reading floor-plans. *Proc. IEEE Int. Conf. on Computer Vision*, 2015. doi: 10.1109/ICCV.2015.255
6. Briechele K., Hanebeck U.D. Localization of a mobile robot using relative bearing measurements. *IEEE Transactions on Robotics and Automation*, 2002, vol. 20, no. 1, pp. 36–44. doi: 10.1109/TRA.2003.820933
7. Thrun S. Probabilistic robotics. *Communications of the ACM*, 2002, vol. 45, no. 3. doi: 10.1145/504729.504754
8. Badrinarayanan V. SegNet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, vol. 39, no. 13, pp. 2481–2495. doi: 10.1109/TPAMI.2016.2644615
9. Shelhamer E., Long J., Darrell T. Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, vol. 39, no. 13, pp. 640–651. doi: 10.1109/TPAMI.2016.2572683
10. Tateno K., Tombari F., Laina I., Navab N. CNN-SLAM: Real-time dense monocular SLAM with learned depth prediction. *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 2017. doi: 10.1109/cvpr.2017.695
11. Laina I., Ruppel C., Belagiannis V. Deeper depth prediction with fully convolutional residual networks. *Proc. 4th Int. Conf. on 3D Vision*, 2016. doi: 10.1109/3dv.2016.32
12. Kendall A., Badrinarayanan V., Cipolla R. Bayesian SegNet: model uncertainty in deep convolutional encoder-decoder architectures for scene understanding. *Proceedings of the British Machine Vision Conference*, 2017. doi: 10.5244/c.31.57
13. Xiao J., Owens A., Torralba A. SUN3D: A database of big spaces reconstructed using SfM and object labels. *Proc. IEEE Int. Conf. on Computer Vision*, 2013. doi: 10.1109/ICCV.2013.458

14. Blanco J.L. Optimal filtering for non-parametric observation models: applications to localization and SLAM // *The International Journal of Robotics Research*. 2010. V. 29. N 14. P. 1726–1742. doi: 10.1177/0278364910364165
15. Sturm J., Engelhard N., Endres F., Burgard W., Cremers D. A benchmark for the evaluation of RGB-D SLAM systems // *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*. 2012. doi: 10.1109/iros.2012.6385773
16. Labbe M., Michaud F. Online global loop closure detection for large-scale multi-session graph-based SLAM // *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*. 2014. doi: 10.1109/iros.2014.6942926
17. Horn B.K.P. Closed-form solution of absolute orientation using unit quaternions // *JOSA A*. 1987. V. 4. N 4. P. 629. doi: 10.1364/JOSA.4.000629
14. Blanco J.L. Optimal filtering for non-parametric observation models: applications to localization and SLAM. *The International Journal of Robotics Research*, 2010, vol. 29, no. 14, pp. 1726–1742. doi: 10.1177/0278364910364165
15. Sturm J., Engelhard N., Endres F., Burgard W., Cremers D. A benchmark for the evaluation of RGB-D SLAM systems. *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2012. doi: 10.1109/iros.2012.6385773
16. Labbe M., Michaud F. Online global loop closure detection for large-scale multi-session graph-based SLAM. *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2014. doi: 10.1109/iros.2014.6942926
17. Horn B.K.P. Closed-form solution of absolute orientation using unit quaternions. *JOSA A*, 1987, vol. 4, no. 4, p. 629. doi: 10.1364/JOSA.4.000629

Авторы

Евстигнеев Максим Игоревич — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, ORCID ID: 0000-0003-1450-0762, mevstigneev@corp.ifmo.ru

Литвинов Юрий Володарович — кандидат технических наук, доцент, старший преподаватель, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, ORCID ID: 0000-0002-8998-7993, yurl13@yandex.ru

Мазулина Вероника Васильевна — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация; аспирант, Технический университет Эйндховена, Эйндховен, 5600MB, Нидерланды, Scopus ID: 57192555805, ORCID ID: 0000-0001-8109-8334, nikasummit@rambler.ru

Authors

Maxim I. Evstigneev — postgraduate, ITMO University, Saint Petersburg, 197101, Russian Federation, ORCID ID: 0000-0003-1450-0762, mevstigneev@corp.ifmo.ru

Yury V. Litvinov — PhD, Associate Professor, Senior lecturer, ITMO University, Saint Petersburg, 197101, Russian Federation, ORCID ID: 0000-0002-8998-7993, yurl13@yandex.ru

Veronika V. Mazulina — postgraduate, ITMO University, Saint Petersburg, 197101, Russian Federation; postgraduate, Technical University of Eindhoven, Eindhoven, 5600MB, Netherlands, Scopus ID: 57192555805, ORCID ID: 0000-0001-8109-8334, nikasummit@rambler.ru