

М. МАНСУР, П. ДАВИДСОН, О. А. СТЕПАНОВ, Ю.-П. РАУНИО, М. М. АРЕФ, Р. ПИШЕ

ОПРЕДЕЛЕНИЕ ДАЛЬНОСТИ НА ОСНОВЕ ДАННЫХ О СОБСТВЕННОМ ДВИЖЕНИИ, ПОЛУЧЕННЫХ С ПОМОЩЬЮ МОНОКУЛЯРНОЙ КАМЕРЫ

Предложен метод определения дальности до объектов, основанный на взаимодополняющем характере последовательных монокулярных изображений и кинематических параметров камеры. Комплексование измерений, полученных с помощью камеры, и кинематических параметров, которые определяются посредством инерциального измерительного модуля (ИИМ) и одометра, осуществляется с помощью обобщенного фильтра Калмана (ОФК). Результаты экспериментов с использованием колесного робота подтвердили результаты моделирования в части ожидаемой оценки точности определения дальности. Показано, что на эффективность предлагаемого метода значительное влияние оказывают взаимное расположение камеры и наблюдаемого объекта, точность измерения параметров движения камеры и пройденное им расстояние. Установлено, что при благоприятных условиях погрешность оценки дальности может составлять всего 1% от расстояния до ориентира. Метод может быть использован для определения дальности до объектов, находящихся в нескольких сотнях метров от камеры.

Ключевые слова: техническое зрение, определение дальности по параметрам движения, обобщенный фильтр Калмана, последовательные изображения, инерциальные датчики, комплексование измерений датчиков.

1. Введение

Последовательность изображений, полученных с помощью движущейся монокулярной камеры в статичной обстановке, содержит подробные сведения как о движении камеры, так и о взаимном расположении и глубине объектов сцены. Это явление известно как эффект параллакса движения, при котором используются точки, соответствующие друг другу на ряде изображений [1]. При наличии изображения трехмерной точки в одной или нескольких проекциях ее местоположение можно восстановить методом триангуляции. Методы восстановления трехмерной сцены по двумерным изображениям с помощью эффекта параллакса в англоязычной литературе известны как «структура из движения»

Мансур Мостафа. Аспирант, Университет Тампере (Финляндия), Университет ИТМО (С.-Петербург).

Давидсон Павел. Доктор наук, старший научный сотрудник, Университет Тампере.

Степанов Олег Андреевич. Доктор технических наук, профессор, начальник научно-образовательного центра, АО «Концерн «ЦНИИ «Электроприбор», Университет ИТМО (С.-Петербург). Вице-президент международной общественной организации «Академия навигации и управления движением».

Раунио Юкка-Пекка. Старший инженер по НИОКР, ААС Technologies (Тампере, Финляндия).

Ареф Мохаммад М. Постдокторант, Университет Тампере.

Пише Роберт. Профессор, Университет Тампере.

(*structure from motion, SFM*). Они основаны на восстановлении сцены по набору изображений на нескольких кадрах и использовании информации о движении камеры. Похожим на SFM способом является определение относительного местоположения и дальности с помощью монокулярного зрения.

Методы SFM, основанные только на изображениях, чувствительны к резким отличиям в оценках соответствия и обычно недостаточно надежны для применения во многих практических областях. Затруднения связаны с неоднозначностью локальной фотометрической информации из-за изменений освещения, наличия препятствий, повторяющихся структур и объектов, движущихся независимо от сцены, и т.д. [2]. Знание параметров движения камеры позволяет повысить эффективность алгоритмов SFM при наличии ее резких перемещений. Комплексирование визуальной информации и оценок движения камеры с помощью инерциальных датчиков все чаще применяется на практике в таких областях, как дополненная или виртуальная реальность, когда визуальную и инерциальную информацию объединяют для оценивания трехмерного углового и поступательного движений камеры [3–5].

Движение проекции объекта на плоскости изображения относится исключительно к кинематическим параметрам камеры: угловой и поступательной скоростям. Уравнения, связывающие координаты и скорость проекции точки-ориентира с поступательными и угловыми скоростями камеры, были выведены Лонге–Хиггинсом и др. [6]. Из них следует, что по монокулярному видению жесткой фактурной поверхности можно отделить движение глаза относительно нее от поля скоростей меняющегося ретинального (*от слова «рети́на» – сетчатка*) изображения. Кроме того, обнаружено, что дальность, рассчитанная по параллаксу движения, полностью задается ретинальной скоростью. В работе [7] показано, что информацию о движении камеры, полученную от инерциальных датчиков, можно также использовать для прогнозирования местоположения элемента сцены, при этом значительно сокращаются объемы вычислений.

Методы оценивания дальности по набору последовательных монокулярных изображений были предложены в работах [3] и [4]. Их авторы используют пиксельный алгоритм для оценивания дальности и ее неопределенности в каждом пикселе с последующим пошаговым уточнением полученных оценок с течением времени. Основанный на ФК алгоритм позволяет обрабатывать последовательные изображения с небольшим межкадровым смещением и получать оценку дальности в реальном времени. Для этого необходимо знать перемещение камеры, поскольку в модели прогноза ФК используются текущая карта дальностей и оценки параметров движения камеры, чтобы предсказывать дальности для следующего изображения. В работах [8, 9] авторы предложили пакетный алгоритм и рекуррентный итерационный ОФК для оценивания движения датчиков и структуры сцены с учетом всех измерений, полученных одновременно от камеры, гироскопов и акселерометров. Наряду с обычными проективными камерами в этих работах использовались многонаправленные камеры и широкоугольные объективы.

В работе [10] рассмотрены слабо- и сильносвязанные схемы для одновременного оценивания структуры сцены и движения по данным камеры и инерциальных датчиков. В слабосвязанном подходе блок инерциальных чувствительных элементов и блок SFM работают отдельно с разной скоростью и обмениваются информацией. В сильносвязанном подходе применяется единый ОФК высокого порядка, который объединяет необработанные визуальные данные и данные инерциальных датчиков в едином фильтре, а не использует два отдельных фильтра в каждом блоке. Вектор состояния включает кинематические перемен-

ные, которые описывают движение камеры и трехмерные координаты N точек сцены. Кроме того, вектор состояния может быть дополнен неизвестными параметрами камеры, а также смещением нулей инерциальных датчиков и их масштабными коэффициентами. Вектор состояния может включать свыше 20 компонент, что усложняет процесс вычислений и настройки. Стратегия сочетания данных ИИМ и визуальной информации эффективна при наличии сбоя в получении визуальной информации и позволяет определять относительное местоположение при минимальных требованиях к системе технического зрения.

Визуально-инерциальные методы SFM и определения дальности можно разделить по категориям в зависимости от сферы их применения. Контрольно-измерительные приборы всегда содержат ИИМ и иногда одометр для колесных роботов или доплеровский лаг для подводных роботов. К сферам применения относятся беспилотные летательные аппараты (БЛА) [11, 12] и автономные необитаемые подводные аппараты (АНПА) [13–15], визуальное сервоуправление для роботов-манипуляторов [16–19] и колесных роботов [20–21]. В случае БЛА или АНПА визуально-инерциальный метод SFM позволяет решать проблемы наличия помех и неопределенной динамики.

Метод SFM может также применяться для формирования карты дальностей в режиме реального времени в областях, связанных с дополненной реальностью (ДР) [22]. Одна из главных технически сложных задач, стоящих перед всеми системами ДР, заключается в точной локализации и отслеживании изображений, движущихся относительно окружающих физических объектов. В этих применениях дальность обычно определяется компьютерно-визуальными методами на основе бинокулярной диспаратности и анализа поля освещенности (shading visual cues). Вместе с тем SFM-методы сложно использовать в режиме реального времени, поэтому в некоторых случаях используются лазерные дальнометры. Еще одна проблема состоит в том, что бинокулярная диспаратность и анализ поля освещенности работают только на небольших расстояниях до объектов сцены. Эти трудности можно преодолеть с помощью робастной визуально-инерциальной одометрии [23].

Некоторые исследователи провели натурные эксперименты для определения точности оценивания дальности в рамках своих методик, а также других важных характеристик, в частности времени сходимости. В работе [15] разработан нелинейный наблюдатель пониженного порядка для оценивания дальности, который реализован для подводного робота. В ходе натурных испытаний время сходимости составило 25 с. Авторы [20] проверили свой метод на колесном роботе, при этом точки-ориентиры находились на расстоянии нескольких метров от камеры; время сходимости в результате эксперимента составило около 30 с. В работе [11] определялась дальность с помощью камеры, установленной на БЛА. Время сходимости при оценивании дальности на уровне 1,5 м на основе ОФК с использованием оптического потока составило 25 с. В [24] также использовался ФК с известными положениями камеры для оценивания дальности до объекта с помощью оптического потока. Время сходимости составило 10 с.

Параллакс движения можно совместить с другими источниками данных о дальности: стереоскопическим зрением, эффектом кинетической глубины, затенением и окклюзией. Это дает возможность получить более стабильную оценку расстояния от наблюдателя до объекта. В исследовании Ланди и др. [25] предлагается комплексировать данные от параллакса движения со стереоскопической диспаратностью, минимизируя несоответствие значений дальности по диспаратности и по параллаксу движения.

Предлагаемая работа основана на предыдущей публикации [21], в которой представлен алгоритм и результаты моделирования задачи определения дальности по параметрам собственного движения. Теоретические результаты [21] в настоящей работе нашли подтверждение в ходе экспериментов в реальных условиях с использованием колесного робота, оснащенного монокулярной камерой, одомером и инерциальными датчиками. Существенно, что предложенный метод позволяет определять относительное местоположение при отслеживании лишь одного ориентира. При этом имеется возможность одновременного отслеживания нескольких ориентиров за счет параллельной реализации одного и того же алгоритма.

Статья построена следующим образом. Во втором разделе описаны система и ее математическая модель, которая связывает измерения, получаемые от изображения, с дальностью и кинематическими параметрами камеры. В третьем разделе представлены постановка и алгоритм решения рассматриваемой задачи. Четвертый раздел содержит информацию о реализации алгоритма, результатах моделирования и натурных испытаниях. Наконец, в пятом разделе представлены выводы о применении предложенного метода монокулярного определения дальности с известными параметрами движения камеры, а также даются рекомендации по дальнейшим исследованиям в этой области.

2. Описание и математическая модель системы

Геометрия камеры и контрольной точки

Будем считать, что имеется неголономный мобильный робот, оснащенный одомером, ИИМ и монокулярной камерой. Предполагается, что робот движется в двухмерной плоскости, а установленная на нем камера обнаруживает стационарный ориентир с координатами $P^N = [X^N, Y^N, Z^N]^T$, заданными в географической навигационной системе координат $\{N\}$. Для упрощения считается, что система координат $\{r\}$, связанная с роботом, и система координат, связанная с ИИМ, согласованы с осью x^r по направлению движения платформы робота вперед, ось z^r направлена вверх, а ось y^r формирует правостороннюю систему координат. Камера установлена таким образом, чтобы ось z^c ее системы координат $\{C\}$ была направлена вдоль оптической оси камеры, сонаправленной оси x^r робота; ось y^c направлена вниз, а ось x^c формировала правостороннюю систему координат. Матрица перехода от системы координат робота к системе координат камеры имеет вид R_r^c .

Геометрия расположения ориентира и камеры в навигационной системе координат представлена на рис. 1. В соответствии с геометрией, как поясняется в [13], координаты ориентира можно записать в виде

$$P^N(t) = q^N(t) + R_c^N P^c(t), \quad (1)$$

где R_c^N – матрица перехода от системы координат камеры $\{C\}$ к навигационной системе координат $\{N\}$; P^c – вектор ориентира в системе координат камеры $\{C\}$; q^N – начало системы координат камеры в навигационной системе координат $\{N\}$. Кинематическое соотношение для этой системы выводится путем дифференцирования (1):

$$\dot{P}^N(t) = \dot{q}^N(t) + \dot{R}_c^N P^c(t) + R_c^N \dot{P}^c(t), \quad (2)$$

$$\dot{R}_c^N = R_c^N [\Omega^c(t)]_x, \quad (3)$$

где

$$[\Omega^c(t)]_x \triangleq \begin{bmatrix} 0 & -\omega_z^c(t) & \omega_y^c(t) \\ \omega_z^c(t) & 0 & -\omega_x^c(t) \\ -\omega_y^c(t) & \omega_x^c(t) & 0 \end{bmatrix} \quad (4)$$

представляет собой кососимметричную матрицу вектора угловой скорости камеры $\Omega^c(t) = [\omega_x^c(t), \omega_y^c(t), \omega_z^c(t)]^T$.

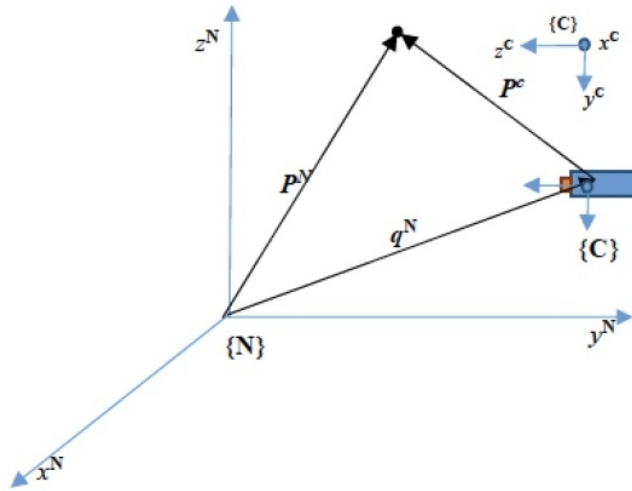


Рис. 1. Геометрия расположения камеры и ориентира в географической системе координат

Принимая во внимание тот факт, что $\dot{P}^N = 0$, и подставив (3) в (2), можно записать:

$$\dot{P}^c = -R_N^c \dot{q}^N(t) - [\Omega^c(t)]_x P^c. \quad (5)$$

Левая часть уравнения (5) описывает кинематику ориентира в системе координат камеры, то есть $\dot{P}^c = [\dot{X}^c, \dot{Y}^c, \dot{Z}^c]^T$, причем первый член правой части уравнения представляет вектор скорости камеры в ее системе координат, а именно:

$$R_N^c \dot{q}^N(t) = [V_x^c(t), V_y^c(t), V_z^c(t)]^T. \quad (6)$$

Для упрощения записи аргумент временной зависимости опускается. Заменив параметры их значениями в (5) и переставив члены уравнения, получаем:

$$\dot{X}^c = -V_x^c + \omega_z^c Y^c - \omega_y^c Z^c, \quad (7a)$$

$$\dot{Y}^c = -V_y^c - \omega_z^c X^c + \omega_x^c Z^c, \quad (7б)$$

$$\dot{Z}^c = -V_z^c + \omega_y^c X^c - \omega_x^c Y^c. \quad (7в)$$

С учетом того что в данном случае рассматривается неголономное двухмерное движение, поворот будет осуществляться только вокруг оси камеры y^c , а вектор скорости будет иметь только одну составляющую в направлении вектора движения. В соответствии с относительным положением камеры в системе координат робота вектор движения будет направлен только по оси камеры z^c . Другими словами, $\omega_z^c = \omega_x^c = 0$ и $V_x^c = V_y^c = 0$. Учитывая эти ограничения модели в (7), получаем следующую кинематическую модель вектора ориентира в системе координат камеры:

$$\dot{X}^c = -\omega_y^c Z^c, \quad (8a)$$

$$\dot{Y}^c = 0, \quad (8б)$$

$$\dot{Z}^c = -V_z^c + \omega_y^c X^c. \quad (8в)$$

Из уравнений (8) видно, что координата Y^c ориентира не изменяется, поскольку робот осуществляет только двухмерное движение.

1) *Модель камеры с малой диафрагмой и формирование изображения.* В работе использовалась монокулярная камера, которую можно описать моделью с малой диафрагмой [1, 26]. В этой модели трехмерный ориентир в системе координат камеры проецируется на точку двухмерного изображения на плоскости изображения камеры (uov) (рис. 2).

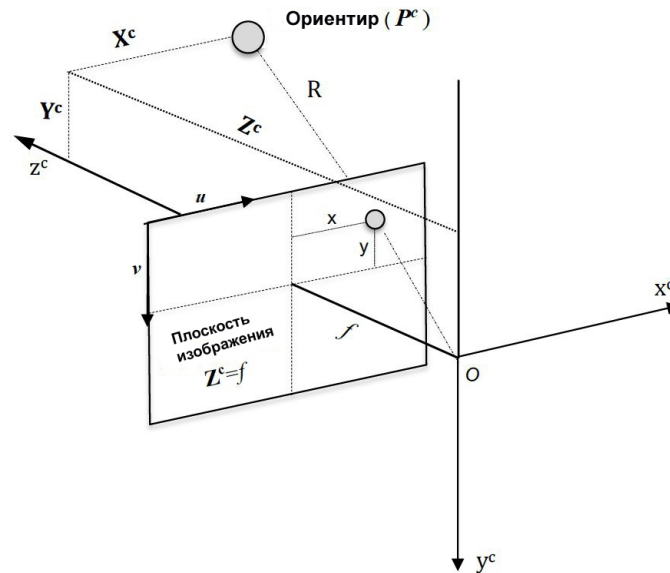


Рис. 2. Перспективная проекция в камере с малой диафрагмой

В соответствии со свойствами перспективной проекции в камере с малой диафрагмой [1, 26] координаты точки изображения можно рассчитать следующим образом:

$$x = f \frac{X^c}{Z^c} + c_x, \quad (9a)$$

$$y = f \frac{Y^c}{Z^c} + c_y, \quad (9б)$$

где (x, y) – координаты точки изображения, f – фокусное расстояние камеры, (c_x, c_y) – координаты «главной» точки. Все указанные параметры измеряются в пикселях. Путем перестановки членов уравнения в (9а) получаем

$$X^c = (x - c_x) \frac{Z^c}{f}. \quad (10)$$

При подстановке (10) в уравнение (8с) имеем следующее уравнение:

$$\dot{Z}^c = -V_z^c + \omega_y^c (x - c_x) \frac{Z^c}{f}. \quad (11)$$

Фокусное расстояние и координаты «главной» точки c_x, c_y являются собственными параметрами камеры. Их необходимо определить до того, как использовать камеру, как будет описано ниже. Кинематические характеристики точки изображения можно получить путем дифференцирования (9):

$$\dot{x} = \frac{Z^c \dot{X}^c - X^c \dot{Z}^c}{(Z^c)^2 / f}, \quad (12a)$$

$$\dot{y} = \frac{Z^c \dot{Y}^c - Y^c \dot{Z}^c}{(Z^c)^2 / f}. \quad (12b)$$

Чтобы избавиться от существенной нелинейности в динамике состояния, координата дальности Z^c в уравнениях (8), (11) и (12) представлена ее обратной величиной $\xi = \frac{1}{Z^c}$ [13]. В работе [27] показано, что благодаря такой параметризации дальности гауссовское распределение достоверно представляет неопределенность дальности в диапазоне от ближайших значений до бесконечности. Подставив (8) в уравнение (12) и применив (11), получаем:

$$\dot{x} = - \left(f + \frac{(x - c_x)^2}{f} \right) \omega_y^c + V_z^c \xi (x - c_x), \quad (13a)$$

$$\dot{y} = V_z^c \xi (y - c_y) - \frac{\omega_y^c (x - c_x)(y - c_y)}{f}, \quad (13b)$$

$$\dot{\xi} = V_z^c \xi^2 - \frac{\omega_y^c \xi (x - c_x)}{f}. \quad (13b)$$

Систему уравнений (13) можно интерпретировать как формирующий фильтр, входными данными которого являются угловая и линейная скорости камеры, а выходными – координаты точки изображения x, y и величина, обратная дальности.

3. Постановка задачи и алгоритм ее решения

Рассмотрим описанный Давидсоном и др. [21, 28, 29] усовершенствованный метод оценивания дальности по последовательности изображений, полученных во время отслеживания ориентира. В работах [21] и [28] авторы представили упрощенную нормированную версию модели (13), где $f = 1$ и $(c_x, c_y) = (0, 0)$. В настоящей работе рассмотрена полная модель камеры, и в (13) используются

собственные параметры камеры, то есть f и (c_x, c_y) . Эти параметры будут описаны в анализе моделирования и определены в процессе калибровки камеры при проведении натурных испытаний.

Постановка задачи

Линию визирования (ЛВ), то есть направление на ориентир, можно определить с помощью камеры. ЛВ изменяется по мере движения камеры, как показано на рис. 3. Изменение ЛВ представляет собой кинематику ориентира и может быть описано уравнением (7).

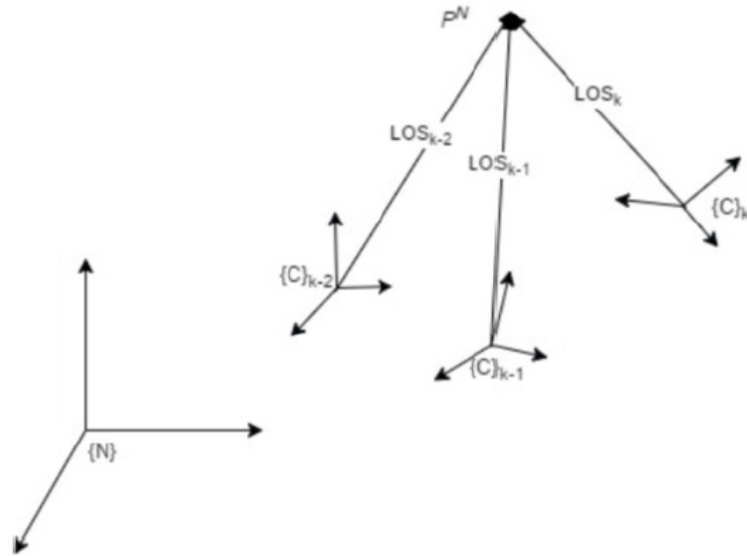


Рис. 3. Формирование линий визирования для моделирования

Соотношение между координатами ориентира и его проекцией на плоскость изображения камеры было описано с помощью системы уравнений (13) с использованием значений линейной и угловой скоростей камеры. Скорости могут быть измерены в системе координат робота, оснащенного одомером и ИИМ, и преобразованы в систему координат камеры с помощью матрицы R_r^c . Поскольку измеренные скорости содержат погрешности, истинные их значения можно представить в виде:

$$\begin{aligned} \omega_y^c &= \tilde{\omega}_y^c - \Delta\tilde{\omega}_y^c, \\ V_z^c &= \tilde{V}_z^c - \Delta\tilde{V}_z^c, \end{aligned} \quad (14)$$

где \tilde{V}_z^c и $\tilde{\omega}_y^c$ – измерения линейной и угловой скоростей камеры, содержащие шумы, а $\Delta\tilde{V}_z^c$ и $\Delta\tilde{\omega}_y^c$ – соответствующие погрешности измерений. Эти погрешности считаются независимыми белыми шумами с функциями автокорреляции $E(\Delta\tilde{V}_{y_i}^c \Delta\tilde{V}_{y_{i\tau}}^c) = \sigma_V^2 \delta(t - \tau)$ и $E(\Delta\tilde{\omega}_{y_i}^c \Delta\tilde{\omega}_{y_{i\tau}}^c) = \sigma_\omega^2 \delta(t - \tau)$ соответственно. При подстановке (14) в (13) система принимает вид:

$$\begin{aligned}\dot{x} &= -\left(f + \frac{(x-c_x)^2}{f}\right)(\tilde{\omega}_y^c - \Delta\tilde{\omega}_y^c) + (\tilde{V}_z^c - \Delta\tilde{V}_z^c)\xi(x-c_x), \\ \dot{y} &= (\tilde{V}_z^c - \Delta\tilde{V}_z^c)\xi(y-c_y) - \frac{(\tilde{\omega}_y^c - \Delta\tilde{\omega}_y^c)(x-c_x)(y-c_y)}{f}, \\ \dot{\xi} &= (\tilde{V}_z^c - \Delta\tilde{V}_z^c)\xi^2 - \frac{(\tilde{\omega}_y^c - \Delta\tilde{\omega}_y^c)\xi(x-c_x)}{f}.\end{aligned}\quad (15)$$

В итоге модель (15) можно записать следующим образом:

$$\dot{X} = \mathbf{S}(X, \tilde{V}_z^c, \tilde{\omega}_y^c, q), \quad (16)$$

где $X = [x, y, \xi]^T$ – вектор состояния, а $q = [\Delta\tilde{\omega}_y^c, \Delta\tilde{V}_z^c]$ – вектор составляющих шумов, которые моделируются в виде нулевых средних значений белых шумов с матрицей спектральных плотностей мощностей

$$Q_c = \begin{bmatrix} \sigma_\omega^2 & 0 \\ 0 & \sigma_V^2 \end{bmatrix}. \quad (17)$$

Начальное значение вектора состояния \hat{X}_0 в уравнении (16) задается в виде нормально распределенного случайного вектора $X_0 \sim N(\hat{X}_0, P_0)$. Начальное приближение может быть получено методом триангуляции, например, с помощью двух (или нескольких) последовательных измерений камеры для грубой оценки расстояния.

В результате измерения на временном интервале k можно представить в виде:

$$\mathbf{z}_k = \begin{bmatrix} x \\ y \end{bmatrix}_k + \mathbf{v}_k \quad (18)$$

или

$$\mathbf{z}_k = HX_k + \mathbf{v}_k, \quad (19)$$

где \mathbf{v}_k – погрешность измерения, представляющая собой дискретный белый шум с матрицей ковариации R , X_k – вектор состояния, H – матрица измерения, определяемая как

$$H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}. \quad (20)$$

В целом рассматриваемая задача сводится к оцениванию обратной величины дальности, задаваемой с помощью формирующего фильтра (16), с использованием полученной с помощью камеры последовательности измерений в виде (18).

Алгоритм решения

В дискретном виде модель (16) можно записать следующим образом:

$$X_k = X_{k-1} + \int_{t_{k-1}}^{t_k} S(X_{k-1}, \tilde{V}_{z_{k-1}}^c, \tilde{\omega}_{y_{k-1}}^c) dt + G(X_{k-1})q_{k-1}, \quad (21)$$

где q_{k-1} – вектор дискретизированных центрированных гауссовских порождающих шумов. Интеграл в (21) можно вычислить различными методами, например методом Эйлера или Рунге–Кутта. В нашем случае применяется метод Рунге–Кутта четвертого порядка. Уравнение (21) можно записать в виде

$$X_k = \psi(X_{k-1}, \tilde{V}_{z_{k-1}}^c, \tilde{\omega}_{y_{k-1}}^c, q_{k-1}). \quad (22)$$

Рассматриваемая задача является нелинейной ввиду нелинейности модели процесса, и для ее решения можно использовать нелинейные алгоритмы байесовской фильтрации. В настоящей работе для этого использован ОФК.

Обобщенный ФК

ОФК представляет собой приближенный вариант оптимального фильтра для задач нелинейной фильтрации [30]. Для прогноза и обновления данных в алгоритме ОФК (первого порядка) выполняются следующие шаги.

Прогноз:

$$X_k^- = \psi(\hat{X}_{k-1}, \tilde{V}_{z_{k-1}}^c, \tilde{\omega}_{y_{k-1}}^c, 0), \quad (23)$$

$$P_k^- = A_{k-1} P_{k-1} A_{k-1}^T + G_{k-1} \frac{Q_c}{\Delta t} G_{k-1}^T. \quad (24)$$

Обновление:

$$K_k = P_k^- H (H P_k^- H^T + R)^{-1}, \quad (25)$$

$$\hat{X}_k = X_k^- + K_k (z_k - H X_k^-), \quad (26)$$

$$P_k = (I - K_k H) P_k^-, \quad (27)$$

где

$$A_{k-1} = \frac{\partial}{\partial X} \psi(X_{k-1}, \tilde{V}_{z_{k-1}}^c, \tilde{\omega}_{y_{k-1}}^c, q_{k-1})_{X=\hat{X}_{k-1}, q_{k-1}=0}, \quad (28)$$

$$G_{k-1} = \frac{\partial}{\partial q} \psi(X_{k-1}, \tilde{V}_{z_{k-1}}^c, \tilde{\omega}_{y_{k-1}}^c, q_{k-1})_{X=\hat{X}_{k-1}, q_{k-1}=0}. \quad (29)$$

Нелинейная модель (22) используется для прогноза состояния в соответствии с (23). Затем модель линеаризуют, чтобы получить матрицы A_{k-1} и G_{k-1} (см. Приложение), которые используются для вычисления матрицы ковариации ошибок прогноза в (24). Обновление измерений (25–27) выполняется так же, как в стандартном ФК.

4. Реализация алгоритма и полученные результаты

Опишем процедуру моделирования алгоритма и результаты его натурных испытаний.

Моделирование

На рис. 4 показана схема получения данных для моделирования. Схема содержит три основных подблока: блок моделирования движения камеры, блок обновления данных о местоположении ориентира и блок моделирования измерений.



Рис. 4. Выработка данных для моделирования

Далее приводится описание каждого из блоков.

Моделирование движения. В этом блоке вырабатываются истинные значения угловой и линейной скоростей камеры. Затем к полученным значениям скоростей добавляются белые шумы для моделирования зашумленных сигналов от ИИМ и одометра.

Обновление координат ориентира. Поскольку камера находится в движении, координаты ориентира в системе координат камеры $\{C\}$ изменяются в каждый момент времени. С помощью приближенного интегрирования (8) по времени можно получить обновление координат ориентира:

$$\begin{aligned} X_k^c &= X_{k-1}^c - \tilde{\omega}_{y_{k-1}}^c Z_{k-1}^c \Delta t, \\ Y_k^c &= Y_{k-1}^c, \\ Z_k^c &= Z_{k-1}^c - \tilde{V}_{z_{k-1}}^c \Delta t + \tilde{\omega}_{y_{k-1}}^c X_{k-1}^c \Delta t. \end{aligned} \tag{30}$$

Расчет точки изображения. Проекцию обновленного ориентира, то есть обновленную точку изображения, можно рассчитать по формуле (9). После этого для моделирования измерений камеры к координатам точки изображения добавляются бел шумные погрешности измерений.

Результаты моделирования

Полученные путем моделирования параметры движения и измерений вводятся в алгоритм на основе ОФК для рекуррентного оценивания вектора состояния X , как было описано в предыдущих разделах. На рис. 5 представлена схема реализации алгоритма ОФК.

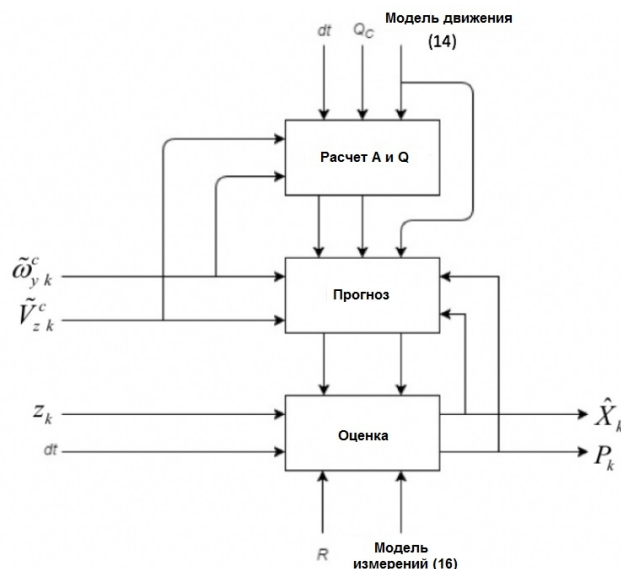


Рис. 5. Реализация алгоритма обобщенного ФК

Параметры, используемые при моделировании, представлены в табл. 1.

Таблица 1

Параметры моделирования

Длительность временного интервала (dt)	0,1 с
Время наблюдения (T)	3 с
σ_ω	0,001 рад/с/ $\sqrt{H_z}$
σ_V	0,01 м/с/ $\sqrt{H_z}$
P_0	диаг. [10 пикс. ² ; 10 пикс. ² ; (3 м) ⁻²]

Обсудим далее зависимость точности оценивания дальности от местоположения ориентира по отношению к фокусу расширения, линейной скорости камеры и так называемую масштабируемость задачи.

Местоположение ориентира (траектория камеры). Фокусом расширения называется точка, в которой отсутствует оптический поток. Эта точка находится на оптической оси камеры. Если ориентир располагается в фокусе расширения или рядом с ним, он будет проецироваться на основную точку камеры или на точку рядом с ней, и, следовательно, между точками изображения на последовательных кадрах практически не будет диспаратности. Таким образом, расстояние до ориентира достаточно сложно оценить, если он находится в фокусе расширения или рядом с ним. В настоящей работе оценивание дальности производилось

по четырем ориентирам. Моделирование осуществлялось методом Монте-Карло; при этом было получено 1000 реализаций погрешностей измерений в соответствии с данными табл. 1 при линейной скорости робота $V_z^c = 0,5$ м/с. Заданная начальная погрешность оценки расстояния составляла 3 м. Ориентиры находились на разном расстоянии от оптической оси, то есть от фокуса расширения.

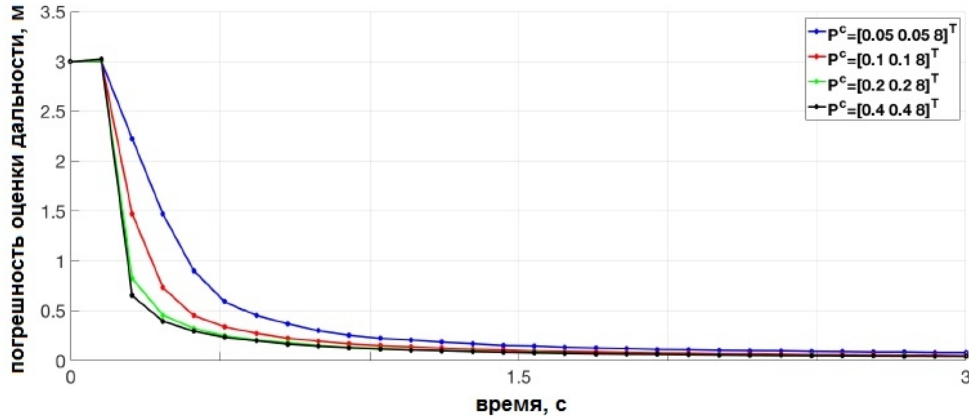


Рис. 6. Влияние расстояния между ориентиром и фокусом расширения на точность оценивания дальности. Моделирование методом Монте-Карло (1000 реализаций) для различных ориентиров

На рис. 6 показан график зависимости точности оценивания расстояния от взаимной геометрии ориентира и камеры. Как видно из рисунка, чем дальше находится ориентир от фокуса расширения, тем быстрее сходится фильтр, что необходимо учитывать при использовании предлагаемого алгоритма. Это означает, что с целью повышения точности скорости сходимости алгоритма траекторию камеры, то есть движение робота, целесообразно задавать так, чтобы ориентиры все время находились на некотором расстоянии от фокуса расширения.

Следует также отметить и тот факт, что при увеличении расстояния между камерой и ориентиром целесообразно увеличивать и расстояние точки изображения от фокуса расширения. Это объясняется тем, что для отдаленных ориентиров диспаратность точек изображения на последовательных кадрах меньше, чем для ближних.

Линейная скорость

Чтобы исследовать влияние линейной скорости камеры на точность оценивания дальности, были проведены испытания для четырех разных значений постоянной линейной скорости камеры: 0,1, 0,3, 0,5 и 0,8 м/с. Параметры, используемые при моделировании, представлены в табл. 1. Ориентир располагался в точке с координатами $P^c = [0,4 \ 0,4 \ 8]^T$.

На рис. 7 показаны средние погрешности оценивания дальности по результатам 1000 циклов моделирования методом Монте-Карло с различными реализациями погрешностей измерений. Можно заметить, что фильтр сходится быстрее при увеличении линейной скорости. Это объясняется тем, что при увеличении скорости диспаратность точек изображения на последовательных кадрах увеличивается, соответственно, угол триангуляции будет больше.

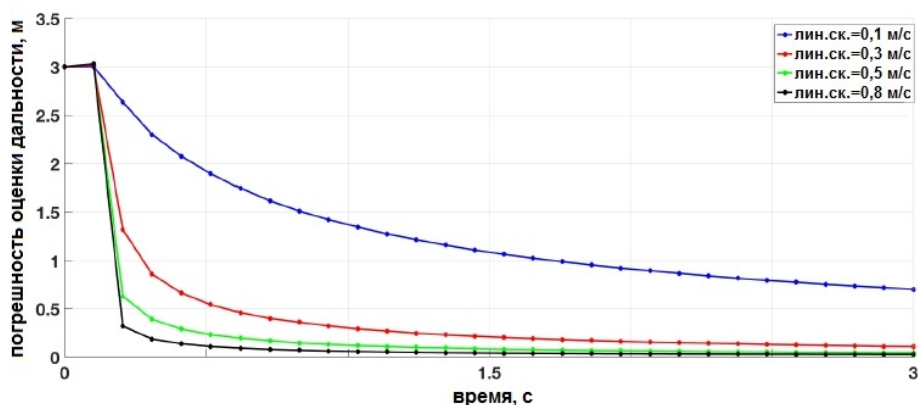


Рис. 7. Зависимость точности оценивания дальности от линейной скорости камеры. Моделирование методом Монте-Карло (1000 реализаций) для различных значений линейной скорости

Масштабируемость задачи

Результаты, представленные на рис. 7, являются масштабируемыми. Оценка дальности зависит не только от линейной скорости камеры, но и от удаленности ориентира. Эту особенность масштабирования можно зафиксировать с помощью параметра $L = VT/D_0$ [32, 33], где V – скорость камеры, усредненная на интервал времени наблюдения; T – продолжительность интервала времени наблюдения; D_0 – начальная дальность до ориентира, т.е. евклидова норма вектора начального положения ориентира в системе координат камеры $\{C\}$ [21, 28].

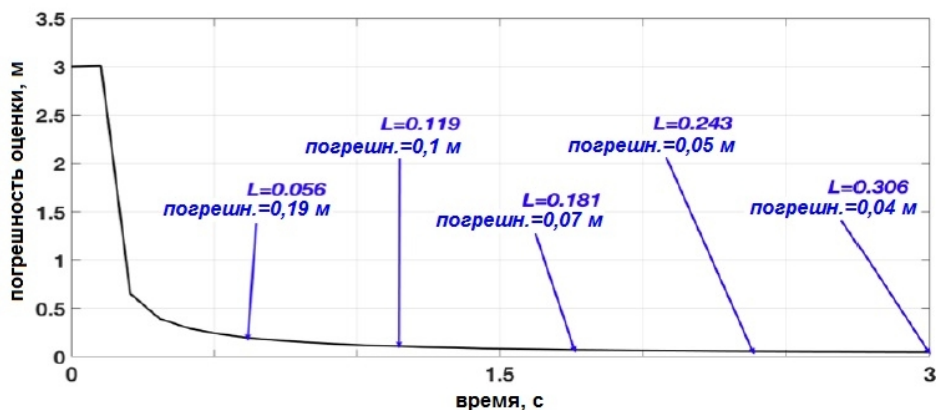


Рис. 8. Зависимость точности оценивания дальности от продолжительности интервала времени наблюдения

Параметр масштабируемости L характеризует отношение пройденного расстояния к начальной дальности до ориентира. Он характеризует способность камеры создавать разницу в видимом положении ориентира, наблюдаемого по двум разным линиям визирования. Следует заметить, что этот параметр не учитывает различные траектории камеры (различные сценарии наблюдения). Поэтому при одном и том же значении L погрешность оценивания дальности будет неодинаковой для разных траекторий. Вместе с тем при одной и той же траектории камеры погрешность оценивания дальности будет одинаковой при тех же значениях L независимо от значения линейной скорости камеры.

На рис. 8 показана зависимость погрешности оценивания дальности от параметра L для камеры, движущейся с линейной скоростью 0,5 м/с и обнаруживающей ориентир с начальными координатами $P^c = [0,4 \ 0,4 \ 8]^T$. Рисунок дает представление об уровне точности, который может быть достигнут при условии, что масштаб (размеры) сцены обзора отличается от масштаба, заданного в представленных экспериментах.

Т а б л и ц а 2

Зависимость точности оценивания дальности от интервала наблюдений

L	0,05	0,12	0,18	0,24	0,31
Относительная погрешность	3,2%	2,1%	1,7%	1,5%	1,35%

Относительную погрешность оценивания дальности можно определить как процентное соотношение погрешности оценивания дальности на любом временном интервале к значению дальности на том же временном интервале для значений параметра масштабирования L , как показано на рис. 8. Можно заметить, что относительная погрешность уменьшается по мере увеличения значения L . В табл. 2 приведены значения относительной погрешности.

Натурные испытания

В этом подразделе представлены результаты натурных испытаний предложенного алгоритма.

Описание оборудования для испытаний. Для испытаний использовался робот *TurtleBot2* фирмы Robotnik (рис. 9).

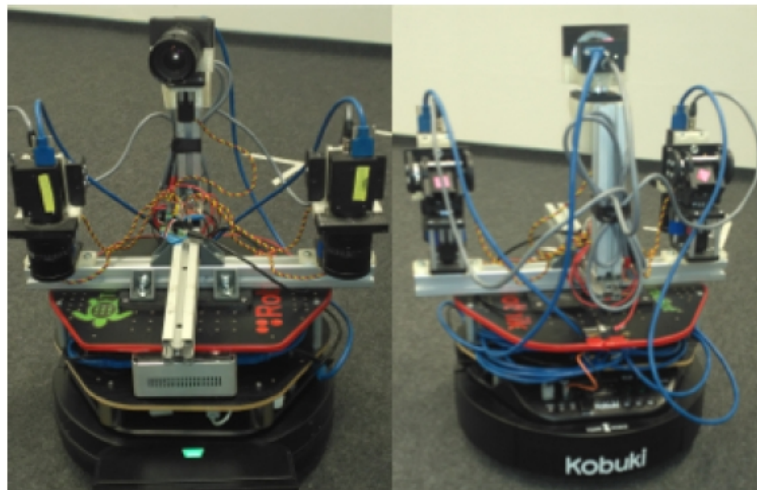


Рис. 9. Робот TurtleBot2. Вид спереди и сбоку.
В эксперименте используется только центральная камера

Платформа робота фирмы Kobuki состоит из откалиброванных ИИМ и одометра и подключена к портативному компьютеру Linux (Ubuntu 14.4). Электро-

питание подается от аккумулятора, установленного в ее нижней части. На работе установлены три камеры Point Grey Grasshopper3, связанные с компьютером. В эксперименте использовалась только центральная камера. На компьютер установлены операционная среда для управления роботами Indigo (ROS Indigo) и драйверы датчиков (ИИМ, одометра и камеры). При запуске робота программные узлы операционной системы позволяют датчикам формировать сообщения с метками времени, которые в удаленном режиме поступают в пользовательский компьютер Linux, соединенный с компьютером робота, через беспроводную локальную сеть LAN. Полученные сообщения сохраняются в одном файле-контейнере операционной системы ROS для последующей камеральной обработки.

Ход испытаний. В режиме ручного управления робот направляется к ориентиру по заранее установленной траектории. В качестве ориентира служит шахматная доска, закрепленная на стене в помещении лаборатории. Робот приводится в действие удаленным пользователем, который передает входные сигналы соответствующего содержания по беспроводной локальной сети. Во время движения робота ИИМ и одометр измеряют его угловую и линейную скорости. Одновременно с этим камера производит съемку ориентира со скоростью 5 кадров в секунду. Сообщения операционной системы с метками времени передаются от ИИМ, одометра и камеры в пользовательский компьютер и сохраняются в файле-контейнере операционной системы. Затем происходит извлечение сигналов из этих сообщений, их синхронизация и обработка.

Предварительная обработка данных и калибровка камеры. Измеренные значения угловой и линейной скоростей преобразуются из системы координат робота в систему координат камеры с помощью матрицы вращения R_r^c . Преобразованные значения скорости сохраняются с соответствующими метками времени. Каждый кадр подвергается обработке для нахождения ориентира.

Калибровка собственной матрицы камеры производится с помощью различных изображений шахматной доски и приложения Camera Calibration Toolbox в среде MatLab [34]. В этом приложении имеется функция *detectCheckerboardPoints* («определение точек шахматной доски»), которая используется для обнаружения шахматной доски. Эта функция строится на алгоритме, предложенном Гейгером и др. в [35]. Этот алгоритм с погрешностью менее пикселя автоматически извлекает углы и совмещает их с (прямоугольными) шахматными досками или клетками. Нас интересует точка изображения в левом верхнем углу шахматной доски. Точки изображения шахматной доски сохраняются с соответствующими метками времени их кадров.

Выявление и отслеживание ориентиров. Полученное с помощью камеры изображение характеризует интенсивность света, отраженного от поверхностей объектов. Интенсивность на изображении может быстро изменяться, например в связи с изменением цвета или ориентации объекта. Точка, в которой интенсивность изменяется в двух направлениях, называется угловой точкой изображения. Существует большое количество различных алгоритмов определения угловых точек. В настоящей работе был применен описанный в [36] алгоритм, который подробно обсуждался в нашей предыдущей работе [21].

Ориентиры можно находить в каждом кадре изображения по отдельности, однако их местоположение неодинаково на разных кадрах, поэтому необходимо

сопоставлять ориентиры на последовательных кадрах, отслеживая их по всему ряду изображений. В наших экспериментах для отслеживания ориентиров применялся алгоритм Канаде–Лукаса–Томази (КЛТ), который представляет собой итеративный метод для пирамид изображений [37, 38]. Трекер КЛТ определяет набор точек изображений, являющихся углами шахматной доски. При этом требуется только одна точка изображения в каждый момент времени, которую нужно отслеживать на последовательных кадрах. Стоит отметить, что для сходимости алгоритма должна быть гарантирована значительная диспаратность точек изображения на двух последовательных кадрах. В ходе экспериментов было установлено, что алгоритм плохо сходится, если диспаратность составляет менее 10 пикселей.

Синхронизация и обработка. Значения скорости и точки изображения сортируются в возрастающем порядке в соответствии с их метками времени. После этого запускается предложенный алгоритм и при этом вектору состояния \hat{X} и матрице ковариации P присваиваются соответствующие начальные значения. Начальное значение времени принимается равным метке времени первого сигнала. Затем происходит обработка сортированных сигналов в следующем порядке. Сначала вычисляется продолжительность интервала времени путем вычитания из текущей метки времени предыдущей метки. Затем сигнал обрабатывается в соответствии с его типом. Например, в случае сигнала угловой скорости вектор состояния изменяется с учетом изменения угловой скорости по формуле (30), причем $\tilde{\omega}_y^c$ принимается равным значению обработанного сигнала, а \tilde{V}_z^c – равным нулю. При наличии сигнала линейной скорости обработка производится при значении \tilde{V}_z^c , равном значению обработанного сигнала, а $\tilde{\omega}_y^c$ – равном нулю. Обновление измерений производится, только если обрабатываемый сигнал является точкой изображения камеры. Предложенный алгоритм был проверен на компьютере с процессором IntelCore i7 3,5 ГГц. Время, требуемое на обновление ОФК по одному ориентиру, составляет около 0,06 мс.

Результаты эксперимента. В ходе эксперимента робот направлялся вперед в сторону ориентира, останавливался, а затем направлялся в обратном направлении. Это движение повторялось, имитируя возможное движение тяжелого робота-подъемника, перемещающего различные грузы по территории склада или по производственному помещению. Такое движение можно охарактеризовать как повторяющееся, с низкой скоростью, малым маневрированием и отсутствием резких поворотов. Контрольное значение дальности было получено с помощью алгоритма *перспективных n-точек*, описанного в [39], с учетом известных размеров шахматной доски.

Т а б л и ц а 3

Соотношение относительной погрешности оценки дальности и параметра масштабирования L на основе фактических данных

L	0,042	0,085	0,127	0,174	0,23
Относительная погрешность	9,05%	2,14%	0,92%	0,26%	0,17%

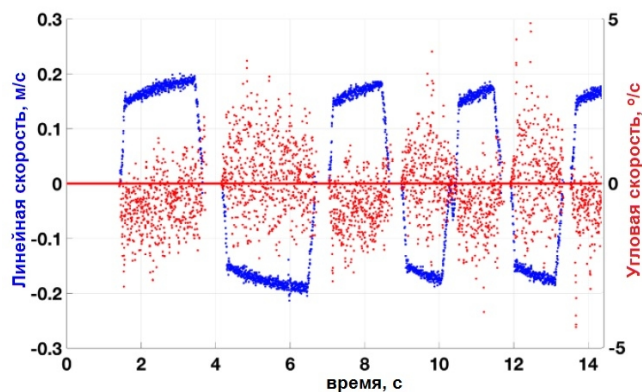


Рис. 10. Линейная (синий цвет) и угловая (красный цвет) скорости камеры

Во время эксперимента робот находился в неподвижном состоянии в течение примерно 1,5 с, а затем начинал серию возвратно-поступательных движений. Значения угловой и линейной скоростей, зафиксированные ИИМ и одомером, представлены на рис. 10. Результаты применения предложенного алгоритма оценивания дальности для реальных данных представлены на рис. 11. Как видно из рисунка, после того как робот начинает движение, фильтр сходится за 1,5-2 с (около 10 кадров). Примерно такое же время требовалось для сходимости фильтра при его моделировании (рис. 8). Расчет параметра масштабирования L выполнялся для движения вперед за первые 20 с. Из рис. 11 можно заметить, что погрешность оценивания дальности уменьшается с ростом значения L . Аналогичный результат был получен при моделировании (рис. 8). Соотношения относительной погрешности и значений параметра L представлены в табл. 3, из которой следует, что при увеличении L относительная погрешность уменьшается, что согласуется с результатами моделирования (табл. 2).

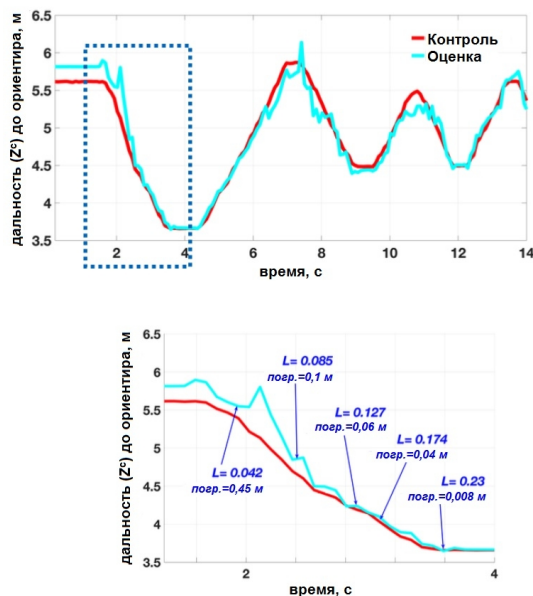


Рис. 11. Оценивание дальности по реальным данным. В нижней части показана сходимость фильтра в течение первых 4 с

На рис. 12 показаны фактические и расчетные значения погрешности при оценивании величины ξ , обратной дальности. Расчет фактической погрешности выполнялся путем нахождения разности между опорным значением и оценкой величины, обратной дальности. Расчетная погрешность представляет собой квадратный корень диагонального элемента матрицы ковариации, определенной с помощью ОФК (σ_ξ). Анализ рис. 12 позволяет сделать вывод об адекватности расчетной характеристики точности, вырабатываемой в фильтре, поскольку фактическая погрешность оценивания находится в пределах интервала погрешности, спрогнозированного с помощью ОФК.

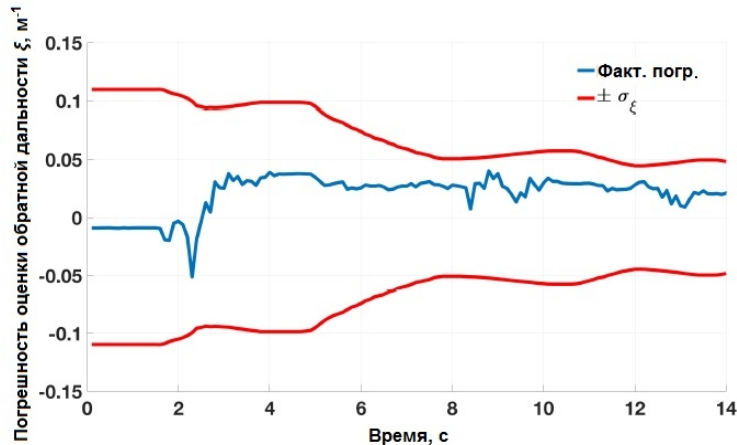


Рис. 12. Фактические и расчетные погрешности оценки величины обратной дальности (ξ)

5. Вывод и дальнейшая работа

В работе предложен метод оценивания дальности с помощью монокулярной камеры, который основан на применении ОФК для комплексирования информации, выделенной из последовательности полученных с помощью камеры изображений, со значениями линейной и угловой скоростей камеры, измеренными соответственно ИИМ и одомером.

Новизна работы заключается в практическом подходе к решению задачи определения дальности по параметрам движения с шестью степенями свободы. Данный подход может быть реализован с помощью мобильного робота, оснащенного стандартной камерой, и математической модели для прогнозирования точности определения дальности в различных сценариях наблюдений.

Предлагаемый метод подходит для применения в режиме реального времени и может быть использован для оценивания расстояний до объектов, удаленных на сотни метров. В основе метода лежит вероятностная модель, учитывающая неопределенность движения камеры. Время сходимости оценки дальности к истинным значениям в несколько раз меньше, чем в других решениях, предложенных в этой области [12]. Кроме того, в работе представлена математическая модель для определения точности алгоритма оценивания дальности для различных сценариев наблюдения, в том числе различных расстояний между камерой и точками, а также различных значений скорости камеры. Ниже обсуждаются некоторые вопросы, связанные с надежностью и областью реализации предложенного метода.

Точность оценивания и время сходимости

Погрешность оценивания дальности зависит главным образом от геометрии взаимного расположения камеры и ориентира. Если при движении камеры образуются сильно отличающиеся виды объекта, оценка расстояния до этого объекта может быть весьма точной – с погрешностью порядка 1% от расстояния. Кроме того, точность зависит от погрешностей определения поступательной и угловой скоростей. Тем не менее в предыдущем исследовании [21] было показано, что для колесных объектов, оснащенных одомером и ИИМ достаточно высокого качества, влияние этих погрешностей на точность оценивания дальности незначительно.

Время сходимости зависит от начального приближения оценки дальности. На примерах, представленных на рис. 6 и 7, начальный вектор состояния был задан с погрешностью дальности 3 м. Уменьшение начальной неопределенности обеспечит ускорение сходимости фильтра. В реальных ситуациях начальное приближение обычно получают с помощью простой триангуляции. В этом случае погрешность оценивания начальной дальности не превышает 50% от расстояния до объекта. Предложенный алгоритм эффективно справляется с такой начальной неопределенностью.

Надежность метода

При обсуждении надежности метода необходимо принять во внимание три важных фактора.

Первый фактор – местоположение ориентира по отношению к фокусу расширения. Как было показано выше, сходимость при использовании предложенного метода происходит быстрее, когда точки находятся далеко от фокуса расширения. Таким образом, чтобы получить быструю сходимость, необходимо задать траекторию таким образом, чтобы ориентир находился на некотором расстоянии от фокуса расширения.

Вторым фактором является характер движения камеры. Поскольку обсуждаемый метод строится на триангуляции, поступательное движение способствует оцениванию дальности: чем выше скорость движения камеры, тем больше изменяется линия визирования, шире угол триангуляции и выше точность. Это подтверждается результатами описанного выше моделирования. В то же время вращение не играет роли в представленном методе, а, напротив, усиливает неопределенность. При вращательном движении линия визирования не изменяется, поэтому угол триангуляции отсутствует и определить дальность с помощью фильтра не представляется возможным.

Третий фактор – начальное состояние фильтра. ОФК использует линеаризованный вариант модели движения. Начальное состояние является исходной точкой линеаризации. Правильный выбор точки линеаризации имеет большое значение, поскольку в противном случае фильтр может расходиться. Существуют, как минимум, три различных способа выбора начального состояния для предложенного метода. *Первый* из них учитывает геометрию пространства, если робот движется в закрытом пространстве. Зная размеры помещения и начальное местоположение робота, можно приблизительно определить начальную дальность до ориентира. Второй способ – использование дополнительного ориентира с известным местоположением в качестве вспомогательной точки. Путем триангуляции двух точек и камеры можно оценить дальность и использовать

полученное значение в качестве начального параметра в ОФК.

Третий способ заключается в получении двух изображений ориентира с двух разных позиций, причем камера в этих двух позициях рассматривается в качестве стереопары. С помощью ИИМ и одометра можно вычислить поворот и переход камеры из одного положения в другое. Зная относительную геометрию камеры в двух положениях, можно оценить приблизительное значение дальности. В настоящей работе применялись первые два способа, поскольку ориентир представлял собой шахматную доску с известной взаимной геометрией углов, а эксперимент проводился в помещении лаборатории.

Область применения

Предложенный метод можно использовать в любых областях, где требуется определить расстояние до ориентира, при выполнении двух условий. Первое из них заключается в том, чтобы движение было преимущественно поступательным, без резких маневров (поворотов). Второе – ориентир не должен находиться слишком близко от фокуса расширения камеры. Одной из наиболее удачных реализаций предложенного метода может быть робот-подъемник на складе или на производственной площадке.

Еще одной важной потенциальной областью применения предложенного метода является составление карт дальностей. Для этого необходимо взять несколько ориентиров (примерно 500-1000) и оценить дальность до каждого из них. Карты дальностей обычно получают с помощью стереокамер, но предложенный метод может стать ключом к использованию с этой целью монокулярной камеры.

Дальнейшая работа

В дальнейшем будет осуществлен сравнительный анализ оценок дальности, полученных с помощью стереозрения и предложенного метода. Последний может быть улучшен за счет использования других нелинейных фильтров, например ансцентного ФК или фильтра частиц, для решения задачи определения дальности по параметрам движения.

Воспользовавшись характерными точками (при их наличии), можно получить значительные преимущества, в частности повышенную точность и надежность. Совместное использование нескольких характерных точек позволяет получить дополнительные измерения, произведенные датчиками, без внедрения лишнего оборудования. Кроме того, этот подход делает более гибкой стратегию использования датчиков, благодаря которой элемент, выпадающий из поля зрения камеры, можно скомпенсировать более подходящим элементом до того, как потеряется первый.

В настоящей работе определяемые ориентиры не связаны с какими-либо конкретными объектами. Текущая задача состоит в получении изображения, на котором каждый выбранный ориентир содержит информацию о дальности или расстоянии до камеры. В дальнейшем работа будет продолжена в направлении разработки методов определения исчезающих точек и ассоциирования данных. Последнее позволит определять конкретные объекты, в том числе движущиеся, и формировать карты окружающей обстановки. Исчезающие точки соответствуют отдаленным ориентирам, следовательно, их проекции на плоскость изображения неподвижны в отсутствие поворотов камеры.

Благодарность

Авторы выражают признательность сотрудникам лаборатории TUT Mobile Robotics Lab за помощь в организации натуральных экспериментов. Работа выполнена при частичной поддержке Российского фонда фундаментальных исследований (проект № 18-08-01101А), а также Правительства Российской Федерации (грант № 08-08).

Приложение

Здесь приведены формулы матриц Якоби (28), (29), используемые в уравнении модели (24).

$$A_{k-1} = \frac{\partial}{\partial X} \psi(X_{k-1}, \tilde{V}_{z_{k-1}}^c, \tilde{\omega}_{y_{k-1}}^c, q_{k-1})_{X=\hat{X}_{k-1}, q_{k-1}=0},$$

$$A_{k-1} = \begin{bmatrix} 1 - \frac{2(x-c_x)\tilde{\omega}_y^c \Delta t}{f} + \tilde{V}_z^c \xi \frac{\Delta t}{f} & 0 & \tilde{V}_z^c (x-c_x) \Delta t \\ -\tilde{\omega}_y^c (y-c_y) \frac{\Delta t}{f} & 1 + \tilde{V}_z^c \xi \Delta t - (x-c_x) \tilde{\omega}_y^c \frac{\Delta t}{f} & \tilde{V}_z^c (y-c_y) \Delta t \\ -\tilde{\omega}_y^c \xi \frac{\Delta t}{f} & 0 & 1 + 2\tilde{V}_z^c \xi \Delta t - (x-c_x) \tilde{\omega}_y^c \frac{\Delta t}{f} \end{bmatrix}_{X=\hat{X}_{k-1}}$$

$$G_{k-1} = \frac{\partial}{\partial X} \psi(X_{k-1}, \tilde{V}_{z_{k-1}}^c, \tilde{\omega}_{y_{k-1}}^c, q_{k-1})_{X=\hat{X}_{k-1}, q_{k-1}=0},$$

$$G_{k-1} = \begin{bmatrix} f + \frac{(x+c_x)^2}{f} & -\xi(x-c_x) \\ \frac{(x-c_x)(y-c_y)}{f} & -\xi(y-c_y) \\ \frac{\xi(x-c_x)}{f} & -\xi^2 \end{bmatrix}_{X=\hat{X}_{k-1}}.$$

ЛИТЕРАТУРА

1. **Hartley, R. and Zisserman, A.**, Multiple View Geometry in Computer Vision. Cambridge University Press, 2003.
2. **Özyesil, O., Voroninski, V., Basri, R., and Singer, A.**, A survey of structure from motion, *Acta Numerica*, 2017, vol. 26, pp. 305–364.
3. **Matthies, L., Szeliski, R., and Kanade, T.**, Incremental estimation of dense depth maps from image sequences, *Proc. Computer Society Conf. on Computer Vision and Pattern Recognition, CVPR'88*, IEEE, 1988, pp. 366–374.
4. **Matthies, L.**, Dynamic stereo vision, *PhD Thesis*, Carnegie Mellon University, 1989.
5. **Hernandez, J., Tsotsos, K., and Soatto, S.**, Observability, identifiability and sensitivity of vision-aided inertial navigation, *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*, 2015, pp. 2319–2325.
6. **Longuet-Higgins, H.C., and Prazdny, K.**, The interpretation of a moving retinal image, *Proc. of the Royal Society London Series B*, 1980, vol. 208, no. 1173, pp. 385–397.
7. **Corke, P.**, An inertial and visual sensing system for a small autonomous helicopter, *Journal of Field Robotics*, 2004, vol. 21, no. 2, pp. 43–51.
8. **Strelow, D. and Singh, S.**, Online motion estimation from image and inertial measurements, *Proc. Workshop on Integration of Vision and Inertial Sensors (INERVIS)*, Coimbra, Portugal, June 2003.

9. **Strelow, D. and Singh, S.**, Motion estimation from image and inertial measurements, *International Journal of Robotics Research*, 2004, vol. 23, no. 12, pp. 1157–1195.
10. **Corke, P., Lobo, J., and Dias, J.**, An introduction to inertial and visual sensing, *International Journal of Robotics Research*, 2007, vol. 26, no. 6, pp. 519–535.
11. **Grabe, V., Bulthoff, H.H., and Giordano, P.R.**, A comparison of scale estimation schemes for a quadrotor UAV based on optical flow and IMU measurements, *Proc. Int. Conf. Intelligent Robots and Systems (IROS)*, IEEE, 2013, pp. 5193–5200.
12. **Delmerico, J. and Scaramuzza, D.**, A benchmark comparison of monocular visual-inertial odometry algorithms for flying robots, *Memory*, 2018, vol. 10, p. 20.
13. **Huster, A. and Rock, S.M.**, Relative position sensing by fusing monocular vision and inertial rate sensors, *Proc. 11th Int. Conf. Advanced Robotics (ICAR'03)*, Coimbra, Portugal, 2003, vol. 3, pp. 1562–1567.
14. **Huster A.**, Relative position sensing by fusing monocular vision and inertial rate sensors, *PhD Thesis*, Stanford University, 2003.
15. **Dani, A., Fischer, N., Kan Z., and Dixon, W.**, Globally exponentially stable observer for vision-based range estimation, *Mechatronics*, 2012, vol. 22, no. 4, pp. 381–389.
16. **Spica, R., Giordano, P.R., and Chaumette, F.**, Plane estimation by active vision from point features and image moments, *Proc. Int. Conf. Robotics and Automation (ICRA)*, 2015, pp. 6003–6010.
17. **Spica, R., Giordano, P.R., and Chaumette, F.**, Active structure from motion: Application to point, sphere, and cylinder, *IEEE Transactions on Robotics*, 2014, vol. 30, no. 6, pp. 1499–1513.
18. **Tahri, O., Boutat, D., and Mezouar, Y.**, Brunovsky's linear form of incremental structure from motion, *IEEE Transactions on Robotics*, 2017, vol. 33, no. 6, pp. 1491–1499.
19. **Martinelli, A.**, Vision and IMU data fusion: Closed-form solutions for attitude, speed, absolute scale, and bias determination, *IEEE Transactions on Robotics*, 2012, vol. 28, no. 1, pp. 44–60.
20. **De Luca, A., Oriolo, G., and Robuffo Giordano, P.**, Feature depth observation for image-based visual servoing: Theory and experiments, *International Journal of Robotics Research*, 2008, vol. 27, no. 10, pp. 1093–1116.
21. **Davidson, P., Raunio, J.-P., and Piché, R.**, Monocular vision-based range estimation supported by proprioceptive motion, *Gyroscopy and Navigation*, 2017, vol. 8, no. 2, pp. 150–158.
22. **Azuma, R.T.**, A survey of augmented reality, *Presence: Teleoperators & Virtual Environments*, 1997, vol. 6, no. 4, pp. 355–385.
23. **Bloesch, M., Omari, S., Hutter, M., and Siegwart, R.**, Robust visual inertial odometry using a direct EKF-based approach, *Proc. Int. Conf. Intelligent Robots and Systems (IROS)*, IEEE, 2015, pp. 298–304.
24. **Matthies, L., Kanade, T., and Szeliski, R.**, Kalman filter-based algorithms for estimating depth from image sequences, *International Journal of Computer Vision*, 1989, vol. 3, no.3, pp. 209–238.
25. **Landy, M.S., Maloney, L.T., Johnston, E.B., and Young, M.**, Measurement and modeling of depth cue combination: In defense of weak fusion, *Vision Research*, 1995, vol. 35, no. 3, pp. 389–412.
26. **Corke, P.**, *Robotics, Vision and Control: Fundamental Algorithms in MATLAB*, vol. 73. Springer Science & Business Media, 2011.
27. **Montiel, J.M., Civera, J., and Davison, A.J.**, Unified inverse depth parametrization for monocular SLAM, *Proc. Robotics: Science and Systems*, 2006.
28. **Davidson, P., Raunio, J.-P., and Piché, R.**, Accurate depth estimation from a sequence of monocular images supported by proprioceptive sensors, *Proc. 23rd Int. Conf. on Integrated Navigation Systems*, St. Petersburg, Russia, 2016, pp. 249–257.
29. **Davidson, P., Mansour, M., Stepanov, O.A., and Piché, R.**, Depth estimation from motion parallax: Experimental evaluation, *Proc. 26th Int. Conf. on Integrated Navigation Systems*, St. Petersburg, Russia, 2019.
30. **Stepanov, O.A.**, Optimal and sub-optimal filtering in integrated navigation systems, in *Aerospace Navigation Systems*; Nebylov, A., Watson, J., Eds., New York: John Wiley & Sons, Inc., 2016, pp. 392–446.
31. **Särkkä, S.**, *Bayesian Filtering and Smoothing*, Cambridge University Press, 2010.
32. **Oshman, Y. and Davidson, P.**, Optimal observer trajectories for passive target localization using bearing-only measurements, *Proc. AIAA Guidance, Navigation, and Control Conference*, San-Diego, CA, 1996, p. 3740.
33. **Oshman, Y. and Davidson, P.**, Optimization of observer trajectories for bearings-only target localization, *IEEE Transactions on Aerospace and Electronic Systems*, 1999, vol. 35, no. 3, pp. 892–902.
34. **Bouguet, J.-Y.**, *Camera Calibration Toolbox for Matlab*, California Institute of Technology, Pasadena, CA, 2006.

35. **Geiger, A., Moosmann, F., Car, O., and Schuster, B.**, Automatic camera and range sensor calibration using a single shot, *Proc. IEEE Int. Conf. Robotics and Automation*, 2012, pp. 3936–3943.
36. **Shi, J. et al.**, Good features to track, *Proc. Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE, 1994, pp. 593–600.
37. **Lucas, B.D. and Kanade, T.**, An iterative image registration technique with an application to stereo vision, *Proc. of Imaging Understanding Workshop*, Vancouver, BC, Canada, 1981, pp. 121–130.
38. **Bouquet, J.-Y.**, Pyramidal implementation of the affine Lucas-Kanade feature tracker: description of the algorithm, *Intel Corporation*, 2001, vol. 5, no. 1-10, p. 4.
39. **Gao, X.-S., Hou, X.-R., Tang, J. and Cheng, H.-F.**, Complete solution classification for the perspective-three-point problem, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, vol. 25, pp. 930–943.

Mansour, M. (ITMO University, St. Petersburg, Russia; Tampere University, Finland), **Davidson, P.** (Tampere University, Finland), **Stepanov, O.A.** (Concern CSRI Elektropribor, JSC; ITMO University, St. Petersburg, Russia), **Raunio, J.-P.** (AAC Technologies, Tampere, Finland), **Aref, M.M., and Piché, R.** (Tampere University, Finland)
Depth estimation with ego-motion assisted monocular camera, *Giroskopiya i Navigatsiya*, 2019, vol. 27, no. 2 (105), pp. 28–51.

Abstract. We propose a method to estimate the distance to objects based on the complementary nature of monocular image sequences and camera kinematic parameters. The fusion of camera measurements with the kinematics parameters that are measured by an IMU and an odometer is performed using an extended Kalman filter. Results of field experiments with a wheeled robot corroborated the results of the simulation study in terms of accuracy of depth estimation. The performance of the approach in depth estimation is strongly affected by the mutual observer and feature point geometry, measurement accuracy of the observer's motion parameters and distance covered by the observer. It was found that under favorable conditions the error in distance estimation can be as small as 1% of the distance to a feature point. This approach can be used to estimate distance to objects located hundreds of meters away from the camera.

Key words: computer vision, depth-from-motion, extended Kalman filter, image sequence, inertial sensing, sensor fusion.

Материал поступил 28.02.2019