

П. ДАВИДСОН, Ю.-П. РАУНИО, Р. ПИШЕ

## ОЦЕНКА ДАЛЬНОСТИ НА ОСНОВЕ МОНОКУЛЯРНОГО ЗРЕНИЯ И ПРОПРИОЦЕПТИВНОГО ДВИЖЕНИЯ

*В статье описывается метод комплексирования измерений монокулярной камеры, движения камеры, а также измерений одометров и инерциальных датчиков скорости. Движение камеры между последовательными изображениями создает базовую линию для расчета дальности путем триангуляции. Рекурсивный алгоритм оценивания основывается на расширенном фильтре Калмана. На точность оценивания глубины сильно влияют взаимная геометрия наблюдателя и ориентира, точность измерения параметров движения наблюдателя и линия визирования на ориентир. С помощью моделирования исследуется, как влияют на точность оценивания погрешности измерения линейной и угловой скорости, шум камеры и траектория наблюдателя. Из результатов этих исследований можно вывести требования к измерительной аппаратуре и сценариям наблюдения. При благоприятных условиях погрешность оценки дальности не превышает 2% от расстояния до ориентира.*

**Ключевые слова:** техническое зрение, определение структуры по отображению движения, ИИМ, гироскоп, одометр.

### Введение

Одна из основных задач технического зрения состоит в восстановлении трехмерной картины окружающего мира на основе двухмерных изображений. Последовательность изображений, получаемых движущейся монокулярной камерой в статичной обстановке, содержит подробные сведения как о движении камеры, так и о взаимном расположении и глубине объектов сцены. Это явление называется эффектом параллакса движения. Восстановление трехмерной сцены с использованием эффекта параллакса обычно называется восстановлением по набору изображений (SFM). В настоящее время ведется активная разработка методов определения взаимного расположения между роботом и окружающими объектами, которые используются при обходе препятствий, картографировании (построении каталога объектов и их положений), локализации (определении положения робота относительно заранее известных объектов) и управлении взаимным расположением робота и объекта для обеспечения наблюдения, моделирования и манипулирования.

**Давидсон** Павел. Доктор наук, научный сотрудник Технологического университета г. Тампере (Финляндия).

**Раунио** Юкка-Пекка. Доктор наук, научный сотрудник Технологического университета г. Тампере.

**Пише** Роберт. Профессор Технологического университета г. Тампере.

Статья по докладу на XXIII Санкт-Петербургской международной конференции по интегрированным навигационным системам, 2016.

Движение камеры между последовательными изображениями создает базовую линию для расчета дальности путем триангуляции. Однако оценивание дальности до выбранных точек часто представляет достаточно большую трудность из-за очень незначительной диспаратности на двух кадрах, особенно если выбранная точка расположена близко к фокусу схода. Эффективные методы обычно предполагают комплексирование данных от камеры с данными от проприоцептивных датчиков, что обеспечивает более надежные измерения по сравнению с алгоритмами, использующими только данные от камеры. Тем не менее комплексирование измерений пеленга от камеры с измерениями собственного движения приводит к появлению нелинейности в задаче оценивания, которую можно решить только с помощью нелинейного алгоритма оценивания, например обобщенного фильтра Калмана. Более эффективные алгоритмы используют специфическое представление вектора состояния, позволяющее уменьшить влияние нелинейности. Эти алгоритмы можно использовать для мобильных роботизированных платформ, оснащенных инерциальными датчиками, одометрами или монокулярными камерами.

Предлагаемый алгоритм обусловлен следующим фактом: люди способны достоверно восстановить трехмерную сцену, не используя бинокулярное зрение (смотря только одним глазом). При этом они полагаются на параллакс движения и собственную вестибулярную систему. В этом случае проприорецепция (ощущение собственного движения в пространстве) поддерживается мозгом, который использует информацию от вестибулярной системы, чтобы понять поступательную и вращательную кинематику тела. Подобный подход, основанный на природных особенностях, можно применить и к техническому зрению с помощью монокулярной камеры и датчиков движения. В данной работе предлагается алгоритм, с помощью которого производится комплексирование данных монокулярной камеры с данными инерциального измерительного модуля (ИИМ) и одометра для оценки взаимного расположения камеры, установленной на мобильном колесном роботе, и стационарного ориентира. Описывается также алгоритм на основе расширенного фильтра Калмана, адаптированный к задаче комплексирования, и представлены результаты моделирования для различных сценариев наблюдения и инструментальных погрешностей. В стратегии измерения используются взаимодополняющие измерения монокулярной камеры, инерциальных датчиков скорости и движение камеры.

### **Аналогичные исследования**

В настоящее время основной метод, используемый для восстановления трехмерных сцен по набору цифровых изображений, – это стереозрение [1, 2]. Стереозрение обычно реализуется с помощью так называемой стереоустановки – двух идентичных камер с параллельными оптическими осями, находящимися на известном расстоянии друг от друга. В данном подходе основная трудность связана с аппаратной стороной: стереоустановка должна быть жесткой, камеры – идентичными, синхронизированными и иметь широкий угол обзора [3]. При быстром движении объекта различия между моментами фиксации кадров недопустимы. Камеры также должны быть тщательно откалиброваны путем расчета поворота и прямолинейного перемещения между ними. Другая задача – совмещение двух изображений от стереоустановки по

точкам, чтобы вывести карты диспаратностей. Это более сложное задание для стереозрения, чем для монокулярного зрения, так как соответствующая точка не всегда находится в том же месте на датчиках изображений в обеих камерах, тогда как в монокулярном зрении соответствующие точки на последовательных изображениях находятся почти в идентичных местах, если частота смены кадров достаточно высока. Используя метод стереозрения, можно точно восстановить глубину расположения объектов ближайшей зоны, но по мере увеличения расстояния точность ухудшается, а для дальних объектов – критично снижается [4].

Трехмерную сцену можно также восстановить, если два изображения, относящиеся к одной области наблюдений, получены последовательно при движении камеры. В том случае когда реконструкция сцены основана на одновременном использовании только одной стереопары, обработка данных происходит с геометрической точки зрения аналогично тому, как это происходит при использовании стереоустановки [2]. Точность можно повысить, если имеется последовательность стереоизображений для реконструкции сцены. Это допускает последовательную обработку стереопар [5]. Данный метод позволяет достичь более высокой точности в силу потенциально широкой полосы пропускания и, соответственно, может успешно использоваться для расчета расстояния до отдаленных объектов.

Когда мы говорим о человеческом зрении, бинокулярная диспаратность – это самый главный источник данных о глубине, если расстояние менее 5 м [4]. Для больших расстояний монокулярные сигналы глубины, особенно параллакс движения, играют важную роль в измерении глубин. Специалисты по психологии восприятия активно изучали параллакс движения и показали, что он имеет первостепенное значение для пространственной ориентации при вождении автомобиля [6] и при посадке самолета [7]. В работе Регана и др. [8] описаны эксперименты, когда пилоты в процессе посадки самолета не могли использовать второй глаз. Качество посадки при этом снизилось незначительно, что говорит о том, что бинокулярные сигналы не столь важны. Многие авторы согласны с Гибсоном [7, 9], что в данных случаях основной упор делается на так называемый фокус схода (или расширения).

В работе Векслера и др. [10] исследуется важность собственного движения для восприятия трехмерной структуры из движущегося изображения. Авторы сравнивали ситуации, когда неподвижный пассивный наблюдатель находился в окружении подвижных жестких объектов и когда активный наблюдатель двигался среди неподвижных объектов. Продемонстрировано, что это не полностью эквивалентная замена, так как, хоть наблюдатели и получали одинаковые визуальные сигналы, восприятие трехмерной структуры у активного наблюдателя зависело от незрительной информации о собственном движении.

В статье Лонге-Хиггинса и др. [11] используется уравнение, связывающее координаты фактурного элемента сцены и его скорость с поступательными и угловыми скоростями наблюдателя. На его основе выводится уравнение, описывающее кинематику ретинального (от слова «рети́на» – сетчатка) положения и скорости. Авторы проанализировали эти уравнения и показали, что по монокулярному видению жесткой фактурной поверхности можно отделить движение глаза относительно нее от поля скоростей меняющегося ретинального изображения. Они также пришли к выводу, что глубина, рассчитанная по параллаксу движения, полностью задается ретинальной скоростью.

Параллакс движения можно совместить с другими источниками данных о глубине: 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. [12] предлагается комплексировать данные от параллакса движения со стереоскопическим рассогласованием, минимизируя разность между глубиной по рассогласованию и глубиной по параллаксу движения.

В работе Ткоча и Янчека [13] разработан метод оценки дальности до ориентира, который можно использовать для камеры, установленной на БЛА. Помимо дальности оценивалась также скорость камеры. Предлагаемый подход основан на применении расширенного фильтра Калмана. Новизна состоит в применении группового алгоритма для инициализации ориентиров. Хотя такой алгоритм способен отслеживать множество точек одновременно, он не дает повышения точности в сравнении с индивидуальным отслеживанием ориентиров. Эффективность алгоритма подтверждается только моделированием при очень оптимистичных допущениях.

Хастлер и Рок [14, 15] предлагают систему, где монокулярное зрение интегрировано с измерениями ИИМ для оценивания дальности от подвижного наблюдателя до стационарного объекта. Подход адаптирован для подводных роботов, которые работают в условиях возмущений и неопределенной динамики. Стратегия комбинирования инерциальных и видеоизмерений устойчива к пропаданию видеоданных и способна определить взаимное расположение при минимальных требованиях к системе технического зрения. Также отмечается, что сочетание ограниченной наблюдаемости и значительной нелинейности, свойственное данной стратегии измерения, создает такую задачу оценки, которую нельзя решить с помощью стандартного обобщенного фильтра Калмана. Чтобы преодолеть эту трудность, была использована особая форма представления состояния, которая обеспечивает линейную модель датчика и переводит все нелинейности в динамику состояния, избегая линеаризаций на шаге коррекции измерения. Прогноз состояния реализуется за счет ансцентной трансформации, которая не требует линеаризации.

Наша работа тесно связана с методом оценки взаимного расположения, предложенным Хастлером и Роком [14, 15], а также с алгоритмом оценки дальности до ориентира Ткоча и Янчека [13]. Алгоритм способен определить взаимное расположение, отслеживая только один ориентир. Однако использование дополнительных элементов, если это возможно, дает значительные преимущества – повышение точности и робастности. Такой мультиэлементный подход обеспечивает наличие дополнительных измерений датчиков без применения добавочной аппаратуры. Кроме того, обеспечивается более высокая гибкость стратегии измерения, таким образом, элемент, выпавший из поля зрения камеры, может быть дополнен более подходящим элементом еще до потери первого элемента.

### **Рекурсивное оценивание глубины по последовательности изображений**

В данном разделе представлен метод оценки глубины (расстояния до ориентира) по последовательности изображений. Направление на ориентир обычно называется линией визирования и измеряется камерой. Геометрия расположения наблюдателя и ориентира показана на рис. 1. Оси  $X_N$ ,  $Y_N$  и  $Z_N$  определяют географическую систему координат  $N$ . Оси  $X_C$ ,  $Y_C$  и  $Z_C$  определяют систему координат камеры  $C$ , где ось  $Z_C$  соответствует оптической оси камеры.

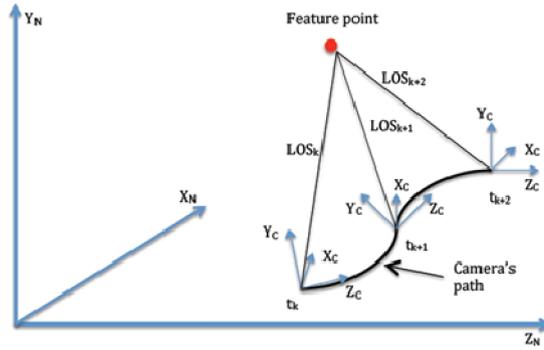


Рис. 1. Взаимная геометрия камеры (наблюдателя) и ориентира

Возможны одновременное отслеживание нескольких точек и оценивание дальности до них с использованием независимых механизмов оценки для каждой точки. Точность оценивания зависит от наблюдаемости точки: точки вблизи фокуса расширения плохо наблюдаемы, следовательно, дальность до этих точек не может быть точно измерена. В будущем можно также рассмотреть потенциальное применение алгоритма координированной оценки. Он позволяет улучшить оценку плохо наблюдаемых точек, налагая дополнительные ограничения. Таким образом, оценка дальности до конкретной точки осуществляется с помощью других точек.

#### Алгоритм оценивания

Алгоритм оценивания построен на основе фильтра Калмана. Используется камера с точечной диафрагмой, которая представлена стандартной моделью в перспективной проекции на рис. 2. Декартовы составляющие проекции ориентира на плоскости изображения обозначены  $x$  и  $y$ . Плоскость изображения обозначена  $Z_C=f$ . Допускается, без нарушения общности, что измерения камеры масштабируются так, что эквивалентное фокусное расстояние равно 1. Ориентир, расположенный в  $R=[X_C, Y_C, Z_C]$  в системе координат камеры, проявляется как элемент в плоскости изображения  $x=X_C/Z_C$  и  $y=Y_C/Z_C$ . Таким образом, измерение камеры  $z$  является проекцией ориентира на плоскость изображения и моделируется следующим образом:

$$z = \begin{bmatrix} x \\ y \end{bmatrix} + n = \frac{1}{Z} \begin{bmatrix} X \\ Y \end{bmatrix} + n. \quad (1)$$

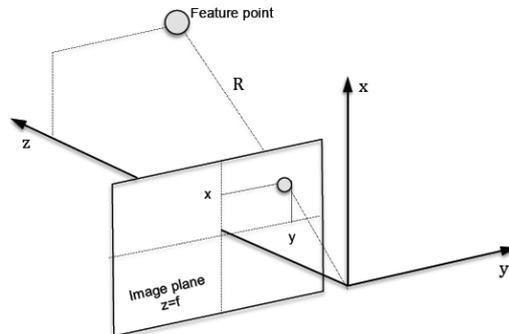


Рис. 2. Модель камеры в перспективной проекции

Погрешности измерений камеры  $n$  считаются центрированным гауссовским шумом. Отслеживание элементов приводит к измерениям расположения элемента на последовательных кадрах. Кинематическое соотношение между ориентиром и его проекцией на плоскость изображения камеры описывается следующим образом [8]:

$$\begin{aligned}\dot{x} &= -v_x \zeta + x v_z \zeta + x y \omega_x - (1 + x^2) \omega_y + y \omega_z \\ \dot{y} &= -v_y \zeta + y v_z \zeta + (1 + y^2) \omega_x - x y \omega_y - x \omega_z \\ \dot{\zeta} &= -v_z \zeta^2 + y \omega_x \zeta - x \omega_y \zeta\end{aligned}\quad (2)$$

В этих уравнениях  $\mathbf{v}=[v_x, v_y, v_z]^T$  – поступательная скорость наблюдателя,  $\boldsymbol{\omega}=[\omega_x, \omega_y, \omega_z]^T$  – угловая скорость. Считается, что эти параметры собственного движения измеряются ИИМ и одометром с достаточно высокой точностью. Мы использовали особую форму представления дальности  $\zeta=1/Z$ , введенную Хастлером и Роком в [9] и [10]. Такое представление снижает доминантные нелинейности в динамике состояния и повышает точность временных коррекций. Алгоритм оценивания построен на основе расширенного фильтра Калмана, где для каждого ориентира измерения и динамика состояния описываются уравнениями (1) и (2) соответственно. При отслеживании нескольких ориентиров необходимо одновременно решать несколько независимых систем уравнений. Однако добавление новых точек не повышает точность оценивания, если для каждой точки используются независимые алгоритмы оценивания.

### Моделирование

Ниже представлены результаты моделирования оценки дальности от камеры до одной точки. Моделирование проводится, чтобы понять, как на точность оценки влияют следующие параметры: погрешности измерения линейной и угловой скорости, шум камеры и взаимная геометрия наблюдателя и ориентира. Во всех случаях считается, что камера движется вперед с постоянной скоростью 0,2 м/с в течение 20 с и преодолевает расстояние 4 м. В большинстве случаев ориентир расположен в  $R=[5, 5, 10]^T$ . В двух случаях с плохой наблюдаемостью ориентира расположены ближе к фокусу расширения в  $R=[1, 1, 10]^T$  и  $R=[0.5, 0.5, 10]^T$  соответственно. Оценка проводится на последовательности из 20 изображений.

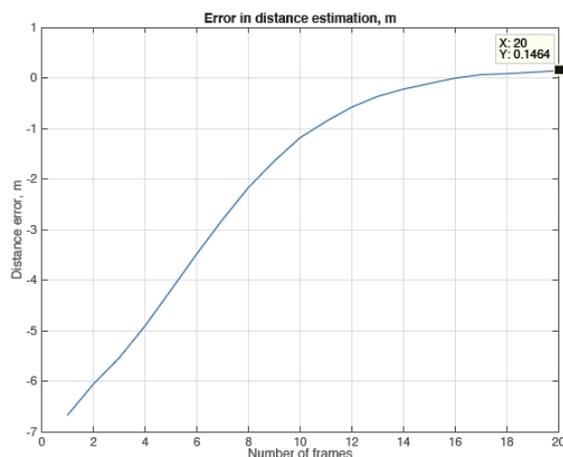


Рис. 3. Погрешность оценки дальности до ориентира при хорошей наблюдаемости без ошибок измерения линейной и угловой скорости. Шум измерений камеры составляет 1 пиксель (0,1 мм)

Влияние траектории наблюдателя

Точность оценивания дальности для случаев хорошей наблюдаемости и отсутствия ошибок одометров и гироскопов показана на рис. 3. Единственный источник погрешностей – это случайный шум в расчетных координатах ориентира на плоскости изображения камеры. Погрешность оценки дальности после 20 изображений составляет около 0,14 м. Схождение фильтра занимает около 15 с.

Точность оценки дальности при плохой наблюдаемости, когда ориентир расположен близко к фокусу расширения  $R=[1, 1, 10]^T$ , показана на рис. 4. В этом случае схождение фильтра происходит медленно, и 20 секунд недостаточно для получения высокой точности. Погрешность оценки дальности после 20 изображений составила около 1,3 м.

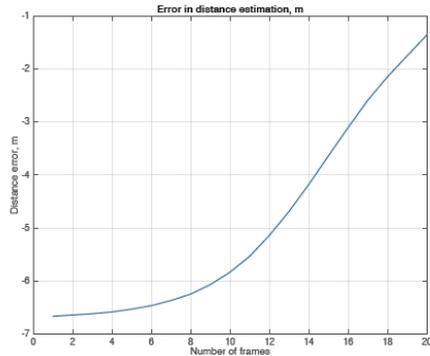


Рис. 4. Погрешность оценки дальности до ориентира при плохой наблюдаемости без погрешностей измерения линейной и угловой скорости. Шум изображения составляет 1 пиксель

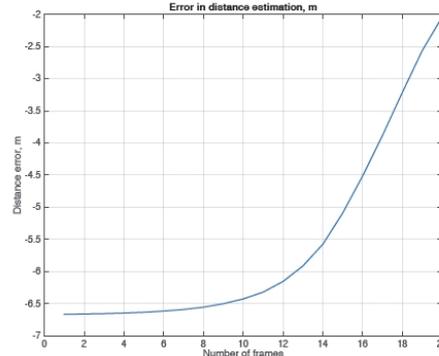


Рис. 5. Погрешность оценки дальности до ориентира при очень плохой наблюдаемости и нахождении точки очень близко к фокусу расширения. Шум изображения составляет 1 пиксель

В следующем примере показана работа фильтра при нахождении ориентира очень близко к фокусу расширения ( $R=[0.5, 0.5, 10]^T$ ). Угол между опорным направлением камеры и линией визирования на ориентир менее 5 град. Результаты оценивания показаны на рис. 5. Схождение происходит очень медленно, и погрешность оценки дальности после 20 изображений – около 2,1 м. По этим результатам можно увидеть, что траектория камеры может сильно влиять на точность оценки дальности. Поскольку данный метод основан на эффекте параллакса движения, траектория камеры во время оценивания должна составлять минимум 30% от дальности до ориентира, а угол между направлением скорости камеры и линией визирования на ориентир должен быть больше 10 град. Пока мы рассматривали только прямолинейные траектории. Использование криволинейной траектории может улучшить точность оценки дальности. Результаты могут размерно варьироваться, и точность оценки глубины зависит от безразмерного параметра, аналогичного тому, что был использован в [16, 17]:

$$K = \frac{VT}{R_0}, \tag{3}$$

где  $V$  – средняя скорость,  $T$  – время наблюдения,  $R_0$  – начальная дальность до ориентира. Параметр  $K$  представляет способность камеры (наблюдателя) создать максимальную разницу в видимом положении ориентира, наблюдаемого по двум разным линиям визирования. Параметр  $K$  не учитывает форму траекто-

рии камеры. Таким образом, при одинаковом  $K$  точность оценивания может быть различной при разных сценариях наблюдения.

### Влияние шума камеры

Данный параметр соответствует шуму измерения камеры, а также случайным ошибкам при слежении за ориентирами. Размер пикселя камеры – 5.5  $\mu\text{м}$ , следовательно, в номинальном случае аддитивный гауссовский шум с СКО 100  $\mu\text{м}$  (0,1 мм) в уравнении (1) будет достоверно представлять такой тип ошибки. Дополнительно были рассмотрены два других случая очень высокого шума измерений 1 мм и 10 мм, чтобы выяснить, когда оценка начинает расходиться. Чтобы исследовать воздействие шума камеры на точность оценивания, рассмотрим случай с хорошей наблюдаемостью без погрешностей линейной и угловой скорости. Номинальный случай показан на рис. 1. Точность оценивания глубины для шума камеры, составляющего 1 мм (в 10 раз больше), показана на рис. 6. Ошибка оценивания дальности после 20 изображений была примерно такой же, как и в случае с номинальным шумом измерения.

При увеличении шума камеры до 10 мм (что невозможно в практических приложениях, данный случай рассматривается, чтобы показать эффект от очень большого шума) схождение фильтра до реальной дальности все же возможно, как показано на рис. 7. Эти результаты демонстрируют, что аддитивный шум измерения не оказывает сколько-нибудь значительного воздействия на точность оценивания координат ориентира на плоскости изображения.

### Влияние погрешностей измерения угловой скорости

Измерение угловой скорости камеры необходимо для расчета координат проекции ориентира на плоскости изображения, как описано в уравнении (2). Этот тип погрешности может быть вызван погрешностями измерения угловой скорости, а также неучтенной погрешностью слежения за ориентацией объекта и дрожанием линии визирования. Данные погрешности могут быть обусловлены, например, вибрациями объекта, когда скорость дискретизации инерциального модуля недостаточно высока, чтобы отслеживать движение объекта. Такой тип погрешности можно моделировать как аддитивную систематическую или случайную погрешность номинальной угловой скорости.

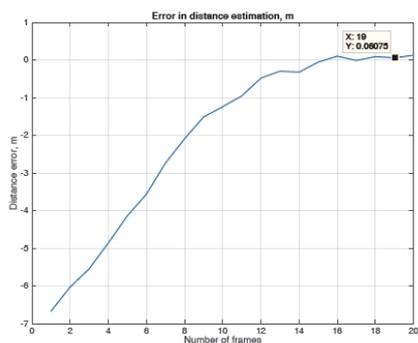


Рис. 6. Погрешность оценки дальности до ориентира при хорошей наблюдаемости и отсутствии погрешностей измерения линейной и угловой скорости. Шум изображения составляет 10 пикселей.

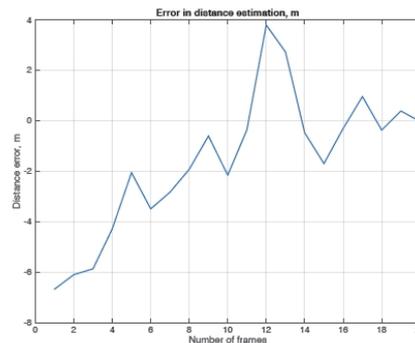


Рис. 7. Погрешность оценки дальности до ориентира при хорошей наблюдаемости и отсутствии погрешностей измерения линейной и угловой скорости. Шум изображения составляет 100 пикселей.

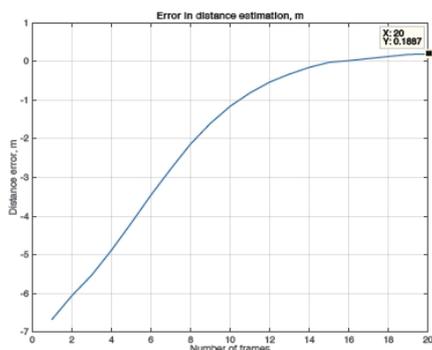


Рис. 8. Погрешность оценки дальности до ориентира при хорошей наблюдаемости и отсутствии погрешностей линейной скорости. Постоянная ошибка гироскопа составляет 10 град/ч. Шум изображения равен 1 пикселю

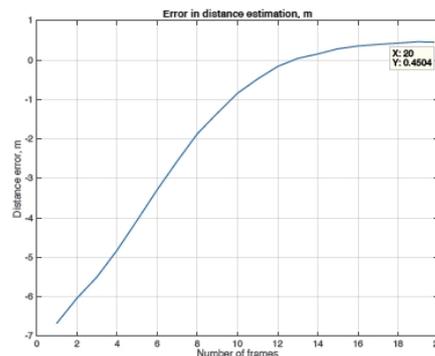


Рис. 9. Погрешность оценки дальности до ориентира при хорошей наблюдаемости и отсутствии погрешностей линейной скорости. Постоянная ошибка гироскопа составляет 100 град/ч. Шум изображения равен 1 пикселю

Чтобы исследовать эффект погрешностей измерения угловой скорости на точность оценивания, предположим, что мы имеем хорошую наблюдаемость и отсутствие погрешностей измерения линейной скорости. Рассмотрим следующие случаи: постоянная ошибка угловой скорости – 10 град/ч, 100 град/ч и 1000 град/ч, случайная аддитивная погрешность угловой скорости с СКО – 1000 град/ч. Точность оценивания глубины при ошибке 10 град/ч показана на рис. 8. Погрешность оценки дальности после 20 изображений составляет около 0,18 м, таким образом, ухудшение точности вследствие малой угловой скорости незначительно. Точность оценивания при ошибке 100 град/ч показана на рис. 9. Погрешность оценки дальности после 20 изображений равна около 0,45 м. Точность оценки при ошибке 1000 град/ч показана на рис. 10. В этом случае схождение фильтра до истинной дальности не произошло. Погрешность оценки дальности после 20 изображений была около 5 м. Такая погрешность измерения значительно выше ограничения, наложенного фильтром. Точность оценки при случайной погрешности угловой скорости 1000 град/ч показана на рис. 11. Оценка дальности зашумлена, но схождение фильтра до истинной дальности происходит. Погрешность оценки дальности после 20 изображений составляет менее 1 м.

#### *Влияние погрешностей измерения линейной скорости*

Чтобы исследовать, как погрешности измерения линейной скорости влияют на точность оценивания, предположим, что мы имеем хорошую наблюдаемость и отсутствие погрешностей измерения угловой скорости. Рассмотрим следующие случаи: постоянная ошибка линейной скорости 0,02 м/с и 0,1 м/с. Точность оценивания глубины при постоянной ошибке по скорости 0,02 м/с показана на рис. 12. Погрешность оценки дальности после 20 изображений составляет около 0,2 м. Следовательно, значительного снижения точности не наблюдается. Точность оценивания при постоянной ошибке по скорости 0,1 м/с показана на рис. 13. Погрешность оценки дальности после 20 изображений составляет около 2 м. Такая погрешность измерения скорости слишком велика, чтобы получить высокую точность оценки дальности.

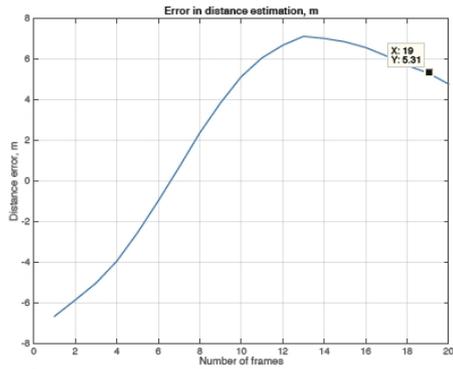


Рис. 10. Погрешность оценки дальности до ориентира при хорошей наблюдаемости и при отсутствии погрешности измерения линейной скорости. Постоянная ошибка гироскопа составляет 1000 град/ч. Шум изображения равен 1 пикселю

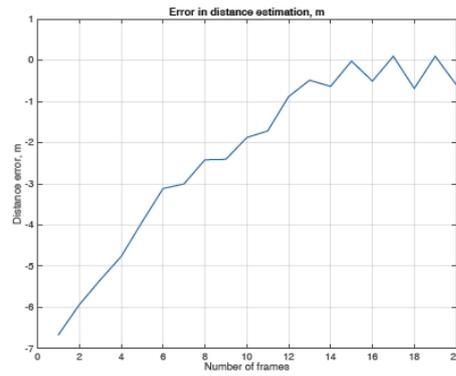


Рис. 11. Погрешность оценки дальности до ориентира при хорошей наблюдаемости и при отсутствии погрешности измерения линейной скорости. Погрешность измерения угловой скорости представляет собой случайный шум с СКО 1000 град/ч. Шум изображения равен 1 пикселю

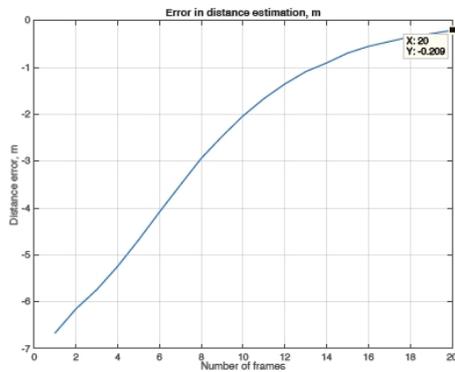


Рис. 12. Погрешность оценки дальности до ориентира при хорошей наблюдаемости и при отсутствии погрешности измерения угловой скорости. Смещение нуля по линейной скорости составляет 0,02 м/с. Шум изображения равен 1 пикселю

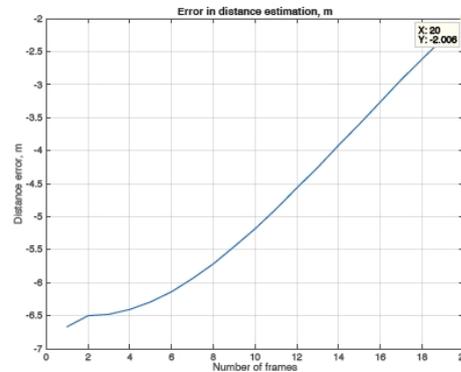


Рис. 13. Погрешность оценки дальности до ориентира при хорошей наблюдаемости и при отсутствии погрешности измерения угловой скорости. Смещение нуля по линейной скорости составляет 0,1 м/с. Шум изображения равен 1 пикселю

### Будущие исследования

Планируется реализовать представленные алгоритмы в двух мобильных роботах, показанных на рис. 14. Первый представляет собой дистанционный манипулятор, установленный на подвижной платформе, а второй – серийную мобильную платформу Robotnik. Оба робота оснащены компьютерами с операционной системой для роботов, инерциальным модулем Microstrain 3DM-GX3-25 и одометром. Эти мобильные платформы также оснащены камерами технического зрения, которые соединяются с компьютером посредством USB-кабелей. Используются следующие датчики: трехосный акселерометр, трехосный гироскоп, трехосный магнетометр, температурные датчики и бортовой процессор, который комплексирует измерения, обеспечивая статические и динамические

измерения ориентации и поступательного движения. Диапазон измерения ориентации и курса в ИИМ составляет 360 град по всем трем осям, частота дискретизации – 100 Гц. Угловая точность ориентации ИИМ – 0,5 град в статических и 2,0 град в динамических условиях. Одометр Renishaw RM22I имеет разрешение 0,7 град и частоту дискретизации 40 Гц. Камеры технического зрения производства Point Grey имеют разрешение 4 Мп и скорость построения изображения 10 Гц. Камера устанавливается на держатели с серводвигателями и допускает наклон до 90 град и панорамирование до  $\pm 90$  град. Поле зрения камер – 48 град.

### Обнаружение ориентиров

В реальных условиях встречаются типичные задачи идентификации хороших визуальных элементов, установления их соответствия и робастного слежения. Алгоритм слежения за элементами отслеживает набор точек по методу Канаде-Лукас-Томаси. Изображение, снятое камерой, описывает интенсивность света, который отображается от объектов. Интенсивность света в изображении может быстро меняться, например, вследствие цветовых изменений или изменения ориентации объекта.

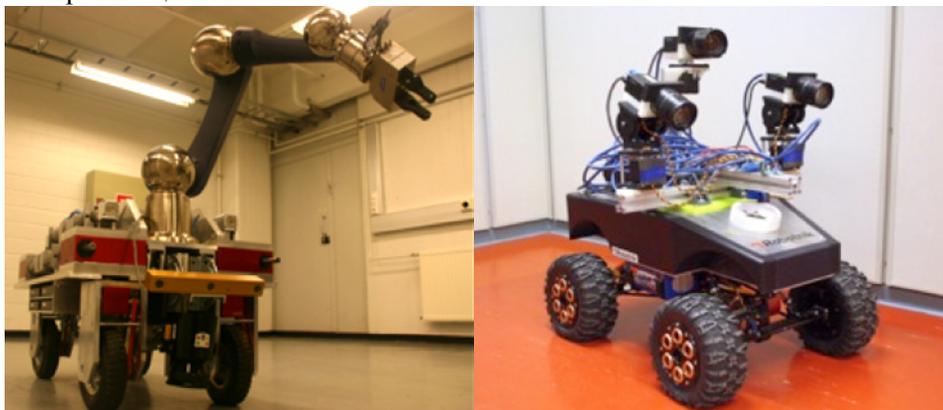


Рис. 14. Дистанционный манипулятор, установленный на подвижной платформе (слева), и серийная подвижная платформа Robotnik (справа)

Если интенсивность меняется одновременно в двух направлениях, такое пятно называется углом изображения. Существует большое многообразие алгоритмов обнаружения углов (см., например, [18, 19]). В данной работе применен алгоритм, описанный в [20]. Обнаружение углов производится по матрице ковариации градиентов, рассчитываемой на площади 25 пикселей ( $5 \times 5$ ). Такая матрица ковариации рассчитывается для положения каждого пикселя в изображении, также рассчитывается минимальное собственное число матрицы, которое служит мерой качества угла. Далее вычисляются местные максимумы для окрестностей  $3 \times 3$ . Наконец, отбраковываются углы, у которых минимальное собственное число меньше меры качества лучшего угла, умноженной на 0,01. Кроме того, алгоритм отбраковывает углы, на расстоянии менее двух пикселей от которых находится более сильный угол. Эти два пороговых предела в алгоритме заданы так, что в каждом изображении находится минимум 500 ориентиров.

В данной работе обнаруженные ориентиры не ассоциированы с каким-либо определенным объектом, и текущая задача заключается в том, чтобы получить изображение, где каждая выбранная точка содержит данные о глубине или

дальности до камеры. В будущей работе мы планируем разработать методы идентификации точек схождения и ассоциации данных. Последнее поможет идентифицировать конкретные объекты, включая подвижные, и построить карту местности. Точки схождения соответствуют дальним ориентирам, соответственно, их проекции на плоскость изображения не двигаются, если камера не поворачивается.

### Отслеживание ориентиров

Ориентиры находятся отдельно по каждому кадру. Однако их положение меняется от кадра к кадру, поэтому элементы в последовательных кадрах нужно соотнести, отслеживая их по последовательности изображений. Метод слежения основан на алгоритме оптического потока, в котором используется итеративный метод Лукаса-Канаде с пирамидами изображений [21, 22]. Метод Лукаса-Канаде допускает, что смещение элементов (т.е. вектор оптического потока) между двумя последовательными изображениями незначительно. Это условие выполняется в наших испытаниях вследствие низкой скорости наземного робота и высокой скорости построения изображения. Оптический поток определяется как вектор  $\mathbf{d}=[d_x, d_y]$ , который минимизирует функцию невязки:

$$e(d_x, d_y) = \sum_{x=u_x-\omega_x}^{u_x+\omega_x} \sum_{y=u_y-\omega_y}^{u_y+\omega_y} (I(x, y) - J(x + d_x, y + d_y))^2, \quad (4)$$

где  $x$  и  $y$  – положения пикселя на плоскости изображения,  $I$  и  $J$  – функции интенсивности последовательных изображений,  $u_x$  и  $u_y$  – координаты точки изображения,  $\omega_x$  и  $\omega_y$  определяют размер области поиска. Минимизация производится для каждого уровня в пирамиде изображений (пирамида – это набор изображений последовательно снижающегося качества по сравнению с оригинальным изображением до достижения некоей заданной крайней точки). В этой работе мы использовали три уровня пирамиды. Минимизация начинается с наименьшего изображения, затем переходит на среднее и заканчивается на оригинальном изображении. Такая пирамидальная методика ускоряет минимизацию и помогает соотнести элементы на различных расстояниях.

### Выводы

В статье описан алгоритм комплексирования измерений монокулярной камеры, инерциальных датчиков скорости и углового движения камеры. Алгоритм обеспечивает определение относительного расстояния между камерой и окружающими объектами и может быть реализован во многих важных практических приложениях, таких как обход препятствий, картографирование, определение положения робота относительно известных объектов и управление взаимным расположением. Моделирование показывает, что данный алгоритм обеспечивает высокую точность оценки глубины (дальности до объектов): погрешность оценки составляет всего 2% дальности до объекта. Однако точность оценки дальности сильно зависит от точности гироскопов и одометров, а также от сценариев наблюдения. При помощи моделирования выработаны точностные требования к аппаратуре и сценариям наблюдения.

Показано, что алгоритм способен работать при достаточно высоких погрешностях измерения угловой скорости (до 100 град/ч). Чтобы получить хорошие

результаты, требуются микромеханические ИИМ средней точности типа Microstrain 3DM-GX3-25 и Xsens MTi-10. Точное измерение линейной скорости также важно, так как оно используется для расчета базовой линии между последовательными кадрами. Поскольку скорость соразмеряется с дальностью до ориентира, погрешность измерения скорости также масштабируется и дается в виде относительной погрешности. По результатам моделирования был сделан вывод о том, что допустимы относительные погрешности измерения скорости до 5% от скорости камеры. Это не очень ограничивающее требование, поскольку обычно колесные датчики положения имеют более высокую точность. Кроме того, следует учитывать скольжение колес.

Взаимная геометрия камеры и ориентира также очень важна. По объективным причинам данный метод не позволяет рассчитать дальность до ориентира, расположенного в фокусе расширения. В силу практических обстоятельств дальность до точек вблизи фокуса схода оценивается с низкой точностью. Моделирование продемонстрировало, что, если угол между ориентирами и фокусом схода составляет более 10 град, значительного ухудшения точности оценки дальности не происходит. Однако если ориентир приближается еще ближе к фокусу схода, точность начинает падать, и для точек, расположенных под углом менее 5 град к фокусу, метод не обеспечивает надежной точности оценки глубины. Если условия приложения допускают некоторую свободу управления траекторией наблюдателя, можно повысить точность, создав такую геометрию между ориентиром и наблюдателем, которая бы максимально увеличивала наблюдаемость системы.

#### ЛИТЕРАТУРА

1. **Lazaros N., Sirakoulis G.C. and Gasteratos A.** Review of stereo vision algorithms: from software to hardware. *International Journal of Optomechatronics*. 2008. 2(4). P. 435–462.
2. **Hartley R., Zisserman A.** *Multiple view geometry in computer vision*. Cambridge university press, 2003.
3. **Vishnyakov B.V., Vizilter Y.V., Knyaz V.A., Malin I.K., Vygolov O.V., and Zheltov S.Y.** Stereo sequences analysis for dynamic scene understanding in a driver assistance system // *SPIE Optical Metrology*. International Society for Optics and Photonics. 2015, June.
4. **Kytö M., Nuutinen M., and Oittinen P.** Method for measuring stereo camera depth accuracy based on stereoscopic vision // *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics. 2011, January.
5. **Beloglazov I.N.** Accuracy of stereoscopic navigation system in flights above natural landscapes and urban land // *Journal of Computer and Systems Sciences International*. 2010. 49.5. P. 802–810.
6. **Gordon D.A.** Static and dynamic visual fields in human space perception // *JOSA*. 1965. Vol. 55. No. 10. P. 1296–1302.
7. **Gibson J.J.** *The ecological approach to visual perception*. Psychology Press, 2013.
8. **Regan D., Beverley K., and Cynader M.** The visual perception of motion in depth // *Scientific American*, 1979.
9. **Gibson J.J.** *The perception of the visual world*. 1950.
10. **Wexler M., Panerai F., Lamouret I., Droulez J.** Self-motion and the perception of stationary objects // *Nature*. 2001. 409(6816). P. 85–88.
11. **Longuet-Higgins H. C. and Prazdny K.** The interpretation of a moving retinal image // *Proceedings of the Royal Society of London. Series B. Biological Sciences*. 1980. Vol. 208. No. 1173. P. 385–397.
12. **Landy M.S. et al.** Measurement and modeling of depth cue combination: in defense of weak fusion // *Vision research*. 1995. 35.3. P. 389–412.
13. **Tkocz M. and Janschek K.** Metric velocity and landmark distance estimation utilizing monocular camera images and IMU data // *Proc. 11th Workshop on Positioning, Navigation and Communication*. IEEE. 2014.

14. **Huster A.** Relative position sensing by fusing monocular vision and inertial rate sensors. Ph.D. dissertation. Citeseer, 2003.
15. **Huster A. and Rock S.M.** Relative position estimation for manipulation tasks by fusing vision and inertial measurements // Proc. 11<sup>th</sup> International Conference on Advanced Robotics. Coimbra, Portugal. 2003, June 30 – July 3. Vol. 2. ICAR. P. 1562–1567.
16. **Hammel S., Liu P., Hilliard E., and Gong K.** Optimal observer motion for localization with bearing measurements // Computers & Mathematics with Applications. 1989. Vol. 18. No. 1. P. 171–180.
17. **Oshman Y. and Davidson P.** Optimization of observer trajectories for bearings only target localization // Aerospace and Electronic Systems, IEEE Transactions on. 1999. Vol. 35. No. 3. P. 892–902.
18. **Harris C. and Stephens M.** A combined corner and edge detector // Proc. of the 4th Alvey Vision Conference. 1988. P. 147–151.
19. **Moravec H.** Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover // Tech Report CMU-RI-TR-3 Carnegie-Mellon University, Robotics Institute. 1980.
20. **Shi J. and Tomasi C.** Good Features to Track // Proc. of the IEEE Conference on Computer Vision and Pattern Recognition. June 1994. P. 593–600.
21. **Lucas B.D. and Kanade T.** An iterative image registration technique with an application to stereo vision // Proc. of Imaging Understanding Workshop. 1981. P. 121–130.
22. **Bouguet J.-Y.** Pyramidal implantation of the Lucas-Kanade Feature Tracker Description of the algorithm. Intel Corporation, Microprocessor Research Labs, 1999.

---

**Davidson P., Raunio J.-P., Piché R.** (Tampere University of Technology, Finland)

Monocular vision based range estimation supported by proprioceptive motion // Гироскопия и Навигация. 2016. Vol. 24. No. 4 (95). P. 98–111.

**Abstract.** This paper describes an approach for fusion of monocular vision measurements, camera motion, odometer and inertial rate sensor measurements. The motion of the camera between successive images generates a baseline for range computations by triangulation. The recursive estimation algorithm is based on extended Kalman filtering. The depth estimation accuracy is strongly affected by the mutual observer and feature point geometry, measurement accuracy of observer motion parameters and line of sight to a feature point. The simulation study investigates how the estimation accuracy is affected by the following parameters: linear and angular velocity measurement errors, camera noise, and observer path. These results impose requirements to the instrumentation and observation scenarios. It was found that under favorable conditions the error in distance estimation does not exceed 2% of the distance to a feature point.

**Key words:** computer vision, structure from motion, IMU, gyroscope, odometer

Материал поступил 07.09.2016