

ОБЗОР СИСТЕМ ВИЗУАЛЬНО-ИНЕРЦИАЛЬНОЙ ОДОМЕТРИИ

Денисов Е. В., магистр, оператор 4 научной роты ФГАУ “Военный инновационный технополис “ЭРА”, г. Анапа, Российская Федерация, email: denisov.evgeni.v@yandex.ru, тел.: 89179378346

Аннотация

В данной работе проведен обзор современных методов в области визуально-инерциальной одометрии, являющейся одной из важнейших тем в современной робототехнике, а также программных систем, реализующий данные методы. Были рассмотрены различные подходы в проектировании систем визуально-инерциальной одометрии в зарубежной литературе, а также были выделены работы с открытым исходным кодом.

Ключевые слова: Военный инновационный технополис “ЭРА”; робототехника; SLAM; визуальная одометрия; инерциальные датчики.

1. Введение

Визуально-инерциальная одометрия — это метод определения позиции с помощью обработки данных с камер и инерциальных датчиков. Системы визуально-инерциальной одометрии (ВИО) являются прямым развитием SLAM систем (от англ. Simultaneous localization And Mapping - Одновременная локализация с построением карты), дополненных использованием инерциальных датчиков для определения точной одометрии, обозначающей нахождение точных координат перемещения платформы с датчиками. Решение проблемы SLAM позволит наделить автономных интеллектуальных роботов точной системой восприятия 3D пространства.

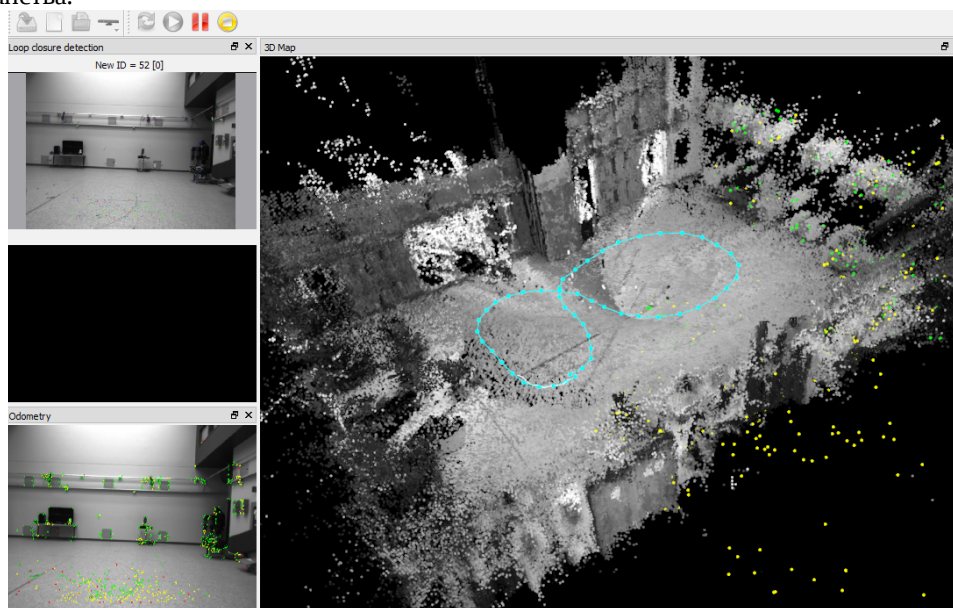


Рис. 1. Пример работы SLAM системы RTAB-MAP [1], определяющей одометрию и реконструирующей заснятое на стереопара 3D пространство.

Системы ВИО состоят из совмещения двух типов одометрии: визуальной и инерциальной, то есть покоящихся на данных с камер и на данных с инерциальных датчиков, (состоящих из гироскопов, акселерометров и других) соответственно. Основная польза совмещения данных типов одометрии состоит в том, что инерциальные датчики показывают более точные данные на коротких промежутках времени, однако имеют тенденцию накапливать ошибку вычисления позы со временем, в то время как камера обладает более широкими возможностями глобальной оптимизации с использованием трехмерного пространства как ориентира для корректировки одометрии во время движения по старым путям.

Данная статья организована следующим образом: секции 2 и 3 описывают два типа одометрии соответственно, в то время как секция 4 описывает работу систем, совмещающих этих двух типов, а секция 5 приводит оценку тенденций развития систем ВИО.

2. Визуальная одометрия

Визуальная одометрия является основой современных SLAM систем благодаря дешевизне камер и возможности трехмерной информации выступать ориентирами при оптимизирующей корректировке реконструкции. Визуальная одометрия состоит из последовательного применения операций из области компьютерного зрения: калибровка камеры (моно, стерео, RGB-D и другие), поиска особых точек (от англ. Features; среди наиболее популярных методов расчета особых точек выделяют ORB, FAST и др.) и вычисления позы камеры в соответствии с изменением положений данных точек, которые вычисляют метод оптического потока, метод сопоставления дескрипторов и другие. Другим важным элементом является оптимизация путей с помощью методов замыкания циклов (от англ. Loop closure) и связанной корректировки (от англ. Bundle adjustment), в основе которых функция минимизации корректного положения точек в 3D пространства и их отражения на всех изображениях (популярно использование библиотек нелинейной оптимизации для C++, таких как g2o, GTSAM и др.). Стоит заметить, что в таких системах используются не все кадры, а только кадры со значительным изменением в позиции съемки. Данные кадры называют ключевыми и их использование снижает вычислительные затраты.

Как и в случае со SLAM системами, процесс одометрии прямо связан с возможности трехмерной реконструкции окружения. Обычно SLAM системам для локализации - а, следовательно, и для одометрии - не нужна реконструкция всего пространства и достаточно участия в локализации лишь особых точек, однако появляются системы с использованием значительно большего числа точек. Тем самым процесс одометрии более простой, чем процесс SLAM.

SLAM системы развиваются уже много лет и для повышения качества работы разработаны различные дополнительные улучшения. DPPTAM [2] заполняет пустые области, получаемые при съемки с моно камеры в следствии монотонности цвета поверхностей окружения, плоскостями. PL-SLAM [3] работает не только с особыми точками, но и линиями между ними, что повышает устойчивость локализации. LSD-SLAM [4] использует в локализации все точки, получаемые операцией градиента на изображении. NID-SLAM [5] способен проводить локализацию в местах с изменяемым освещением. CNN-SLAM [6] выполняет работу с использованием нейронных сетей. DYN-SLAM [7] работает с динамическим окружением. Существуют также работы ([8] и др.), посвященные автономному исследованию пространства роботами на основе SLAM систем. Некоторые системы упрощают модели мира для более эффективной работы в помещениях с полностью ортогональными коридорами (называемых мирами Манхэттена и Атланты [9]). Характерные проблемы с камерами в SLAM системах (плохая работа при низкой или очень высокой освещенности и другие) также в современных работах пытаются решить использованием других сенсоров, в частности вращающимся по двум осям лидаром [10].

Пример схемы процесса визуальной одометрии представлен на Рис. 2. Результаты большинства таких систем доступны в интернете и широко распространяются в удобном для восприятия формате видео в базе видеоданных Youtube.

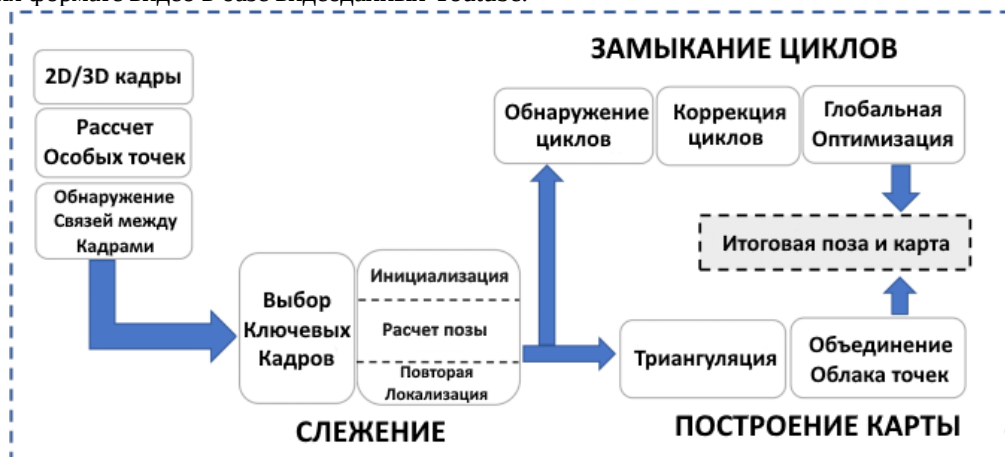


Рис.2. Пример схемы процесса визуальной одометрии

3. Инерциальные одометрия

Инерциальная одометрия в своей основе использует инерциальные устройства, чаще всего состоящие из гироскопа и акселерометра. В области инерциальной одометрии доминирует расширенная нелинейная версия фильтра Калмана [11]. Схема работы фильтра Калмана

представлена на Рис. 3. Несмотря на то, что фильтр Калмана является оптимальным вероятностным алгоритмом для данных, подчиняющихся нормальному распределению, в робототехнике часто необходимо снизить объемы вычислений (скажем, для расчетов на борту небольшого беспилотного летательного аппарата) и в данных случаях предпочтительнее использовать фильтры Маджвика [12], снижающего вычислительную сложность использованием градиентного спуска и часто применяемого при работе с квадрокоптерами, и Шмидта-Калмана [13], снижающего размерность вектора состояния в фильтре Калмана. Главным минусом одометрии с использованием инерциальной одометрии является накопление ошибки со временем, поэтому использование инерциальной одометрии самой по себе не рекомендуется.



Рис. 3. Схема работы фильтра Калмана

4. Совмещение визуальной и инерциальной одометрии

Особенностью совмещения визуальной и инерциальной одометрии является вопрос способа совмещения двух источников данных, на основе которых вычисляются две позиции. В данном вопросе есть два подхода: плотное и неплотное соединение. Плотное соединение предполагает фильтрацию данных, в то время как неплотное предполагает работу двух видов одометрии параллельно без влияния друг на друга до конца приема данных. В соответствии с обзорной статьей [14] плотное соединение за последние несколько лет практически полностью вытеснило неплотное соединение в разрабатываемых системах. В плотном подходе инерциальные датчики чаще всего работают с большей частотой, чем камеры, поэтому данные с инерциальных датчиков обычно совмещают для совпадения с частотой работы камеры. Данный процесс в литературе называют пре-интеграцией инерциальных данных [15].

В соответствии с другими обзорами ([14] и [17]) методы ВИО подразделяются на фильтрационные и оптимизационные. Авторы представленных выше обзоров выделяют фильтрационные методы просты и как не требовательные к вычислительным ресурсам, в то время как оптимизационные обладают наибольшей точностью.

Фильтрационные методы основаны на фильтре Калмана, однако для связи двух типов одометрии был разработан фильтр MSCKF [16], который является стандартом фильтрационных методов. Будучи расширением фильтра Калмана, MSCKF увеличивает количество векторов состояний для запоминания нескольких старых состояний (от англ. Multi-State Constraint Kalman Filter) в рамках подхода сдвигаемого окна (от англ. Sliding Window). Основой данного алгоритма является хранение в памяти последних кадров с последующей оптимизацией одометрии в рамках этих изображений по фильтру Калмана. Оптимизационные методы пришли из SLAM систем, основанных на нелинейной оптимизации вычисляемого пути. Основным отличием оптимизационных подходов от фильтрационных является системы на основе всех изображений на протяжении всего процесса получения данных. Примерами фильтрационных систем являются программы Maplab [18], MSCKF_VIO [19], ACK-MSCKF [20], R-VIO [24] и др. Примерами оптимизационных систем являются VINS-моно [21], OKVIS [22] и др.

Развитие получают системы с заданными ограничениями под конкретными платформами для камер, что особенно полезно при записи с камер, установленных на роботах и транспортных средствах. О них подробнее написано в секции 5.

4.1. Формальное описание системы визуально-инерциальной одометрии

Следующее формальное описание систем ВИО описано в работе [14].

Системы ВИО использует инерциальные данные для прогнозирования (экстраполяция) состояния. Вектор состояния содержит состояние инерционных датчиков x_I и позиции особой точки ${}^G p_f$:

$$\begin{aligned} x &= \begin{bmatrix} x_I^T & {}^G p_f^T \end{bmatrix}^T \\ &= \begin{bmatrix} {}^I_G q^T & b_g^T & {}^G v^T & b_a^T & {}^G p^T & {}^G p_f^T \end{bmatrix} \end{aligned} \quad (1)$$

где ${}^I_G q$ - единичный кватернион, определяющий поворот от глобального фрейма $\{ G \}$ к фрейму инерциальных данных $\{ I \}$; ${}^G p$ и ${}^G v$ - позиция и скорость инерциального датчика в рамках глобального фрейма; b_g и b_a определяют погрешности гироскопа и акселерометра соответственно.

Так как особая точка статична, а инерционные датчики подвижны, непрерывная динамика состояния (1) представляется:

$$\begin{aligned} {}^I_G \dot{\bar{q}}(t) &= \frac{1}{2} \Omega({}^I \omega(t)) {}^I_G \bar{q}(t), {}^G \dot{p}(t) = {}^G v(t), {}^G \dot{v}(t) = {}^G a(t) \\ \dot{b}_g(t) &= n_{wg}(t), b_a(t) = n_{wa}(t), {}^G \dot{p}_f(t) = 0_{3 \times 1} \end{aligned} \quad (2)$$

где ${}^I \omega = [\omega_1 \ \omega_2 \ \omega_3]^T$ - угловая скорость инерциальных датчиков, выраженных в фрейме $\{ I \}$, ${}^G a$ - ускорение инерциальных датчиков в фрейме $\{ G \}$, n_{wg} и n_{wa} - функции белого гауссовского шум, влияющие на погрешности инерциальных датчиков,

$$\Omega({}^I \omega) = \begin{bmatrix} 0 & \omega_3 & -\omega_2 \\ -\omega_3 & 0 & \omega_1 \\ \omega_2 & -\omega_1 & 0 \end{bmatrix}$$

, где $\begin{bmatrix} \omega_1 & \omega_2 & \omega_3 \end{bmatrix}^T$ - кососимметричная матрица. Измерения гироскопа и акселерометра выражены в ω_m и a_m соответственно в локальном фрейме $\{ I \}$ и представлены:

$$\omega_m(t) = {}^I \omega(t) + b_g(t) + n_g(t) \quad (3)$$

$$a_m(t) = C({}^I_G \bar{q}(t)) ({}^G a(t) - {}^G g) + b_a(t) + n_a(t) \quad (4)$$

где ${}^G g$ - гравитационная акселерация, выраженная в $\{ G \}$, а n_g и n_a определяют нулевое среднее значение белого гауссовского шума.

Линеаризация формулы (2) в текущем состоянии производится с помощью непрерывного расчета состояния через модель прогнозирования:

$$\begin{aligned} {}^I_G \hat{\bar{q}}(t) &= \frac{1}{2} \Omega({}^I \hat{\omega}(t)) {}^I_G \hat{\bar{q}}(t), {}^G \hat{\dot{p}}(t) = {}^G \hat{v}(t), {}^G \hat{\dot{v}}(t) = {}^G \hat{a}(t) \\ \hat{\dot{b}}_g(t) &= 0_{3 \times 1}, \hat{\dot{b}}_a(t) = 0_{3 \times 1}, {}^G \hat{\dot{p}}_f(t) = 0_{3 \times 1} \end{aligned} \quad (5)$$

где $\hat{a} = a_m - \hat{b}_a$ и $\hat{\omega} = \omega_m - \hat{b}_g$. Ошибка состояния размерностью 18 x 1 определена следующим образом [см. (1)]:

$$\tilde{x}(t) = \begin{bmatrix} {}^I \tilde{\theta}^T(t) & \tilde{b}_g^T(t) & {}^G \tilde{v}^T(t) & \tilde{b}_a^T(t) & {}^G \tilde{p}^T(t) & {}^G \tilde{p}_f^T(t) \end{bmatrix}^T \quad (6)$$

где приведена умножающая ошибка для кватерниона, то есть ошибка между кватернионом \bar{q} и его вычислением $\hat{\bar{q}}$ как 3 x 1 вектор угловой ошибки ${}^I \tilde{\theta}$, расширенно обозначенная как

$$\delta \bar{q} = \bar{q} \otimes \tilde{\bar{q}} \approx \begin{bmatrix} \frac{1}{2} {}^I \tilde{\theta} \\ 1 \end{bmatrix}$$

ошибка кватернион: $\delta \bar{q}$ описывает небольшой поворот,

$$E \begin{bmatrix} I \tilde{\theta} & I \tilde{\theta}^T \end{bmatrix} \text{), для отношения неопределенного.}$$
$$\dot{\tilde{x}}(t) = F_c(t)\tilde{x}(t) + G_c(t)n(t) \quad (7)$$

(8)

Камера выводит особые точки, которые затем используются для вычисления движения платформы с установленной на ней камерой. В случае откалиброванной камеры, измерение

особой точки на временном шаге k - это проекция 3D точки ${}^C_k p_f$ от перспективы, выраженная в текущем фрейме камеры $\{C_k\}$ на плоскость изображения в виде,

$$z_k = \frac{1}{z_k} \begin{bmatrix} x_k \\ y_k \end{bmatrix} + n_{f_k} \quad (13)$$

$$\begin{bmatrix} x_k \\ y_k \\ z_k \end{bmatrix} = {}^C_k p_f = C({}^C_k q) C({}^C_k q_k) ({}^G p_f - {}^G p_k) + {}^C p_I \quad (14)$$

где n_{f_k} - это измерение белого гауссовского шума с нулевым математическим ожиданием и ковариацией R_k . В (13), $\{C_k, {}^C_k q, {}^C p_I\}$ означают поворот и перемещение между камерой и инерциальными датчиками. Эта трансформация может быть найдена, к примеру, калибровкой внешних параметров камеры перед проведением экспериментов. Однако, на практике идеальная калибровка может быть недоступна и тогда для правильной работы ВИО необходимо включить эти калибровочные параметры в вектор состояния и последовательно рассчитывать их вместе с позами инерциальных датчиков и камеры.

Для использования расширенного фильтра Калмана, линеаризация формулы (13) необходима для вычисления следующего остатка измерения [см. (6)]:

$$z_k = H_k \bar{x}_{k|k+1} + n_{f_k} = H_{I_k} x_{I_{k|k+1}} + H_{f_k} {}^G p_{f_{k|k+1}} + n_{f_k} \quad (15)$$

где измерение Якобиана H_k рассчитывается как:

$$H_k = \begin{bmatrix} H_{I_k} & H_{f_k} \end{bmatrix} \quad (1)$$

$$= H_{proj} C({}^C_k q) \begin{bmatrix} H_{\theta_k} & 0_{3 \times 9} & H_{p_k} & C({}^C_k q_k) \end{bmatrix} \quad (6)$$

$$H_{proj} = \frac{1}{\hat{z}_k^2} \begin{bmatrix} \hat{z}_k & 0 & -\hat{x}_k \\ 0 & \hat{z}_k & -\hat{y}_k \end{bmatrix} \quad (1)$$

$$C({}^C_k q_k) = \begin{bmatrix} \cos q_k & \sin q_k & 0 \\ -\sin q_k & \cos q_k & 0 \\ 0 & 0 & 1 \end{bmatrix}, H_{p_k} = \frac{1}{\hat{z}_k^2} \begin{bmatrix} \hat{z}_k & 0 & -\hat{x}_k \\ 0 & \hat{z}_k & -\hat{y}_k \end{bmatrix} \quad (8)$$

После того, как измерение Якобиана и остатка рассчитаны, мы можем применить стандартное обновление расчёта состояния и ковариации ошибки, используя стандартный расширенный фильтр Калмана.

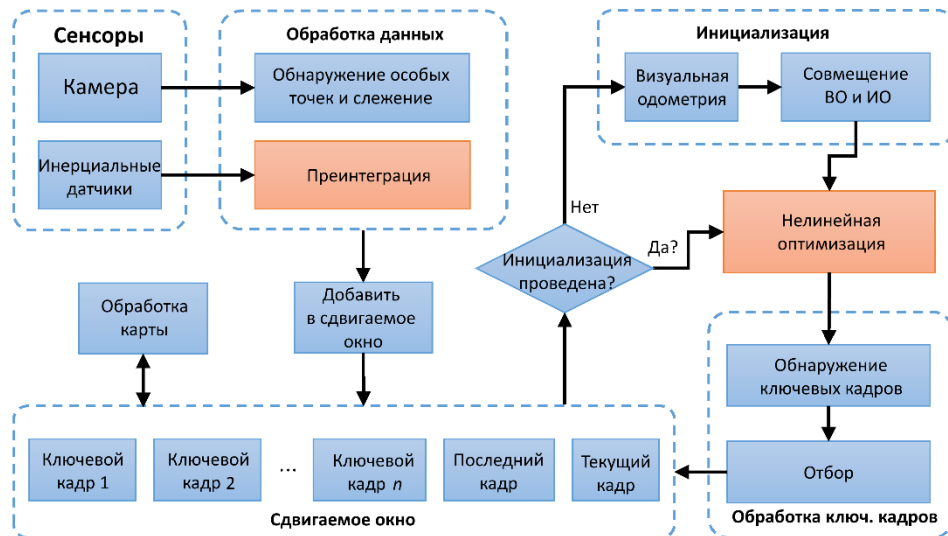


Рис. 4. Пример схемы процесса визуальной-инерциальной одометрии на базе MSCKF [16]

5. Тенденции в области визуально-инерциальной одометрии

Тем не менее, несмотря на большой прогресс области в последние годы и улучшенный результат по сравнению с системами визуальной одометрии, системы визуально-инерциальной одометрии не позволяют достигать точной локализации и не решают основную проблему SLAM систем. Также, как и визуальные SLAM системы, они уязвимы к динамическим изменениям и к большим объемам данных на больших местностях [15]. Другие расширения, представленные в конце секции 2, также подходят для ВИО систем.

Стоит подробнее описать нововведения существующих современных систем. MSCKF_VIO [19] оптимизирует стерео систему до такого уровня, что она способна работать на компьютере невысокой производительности без видеокарты. R-VIO [23] переформулирует математическую основу системы для отказа от использования глобальных координат в фильтрации, что требует больших вычислительных ресурсов, однако это повышает локализацию из-за отсутствия проблемы несоответствия наблюдений. StructVIO [9] работает в условиях мира Атланты, что значительно повышает устойчивость локализации внутри зданий. Многие из представленных систем доступны с открытым исходным кодом, а Matlab [18] позволяет совмещать различные элементы системы для упрощения подбора подходящих алгоритмов для получения наилучших результатов.

Перспективной тенденцией является переход к ограниченным моделям одометрии, преграждающих системам выдавать значения позиции камеры, которые не свойственны платформе движения камеры. Скажем, наземный робот не может оторваться от земли, однако система одометрии может проигнорировать столь очевидный для человека принцип гравитации. ACK-MSCKF [20] ограничивает движение системы до автомобильного движения с рулевым принципом Аккермана, представленный на Рис. 5, тем самым добиваясь лучших результатов в сравнении с системами, не имеющих таких ограничений и предполагающих качественную работу при любой конфигурации платформы (Рис. 6).

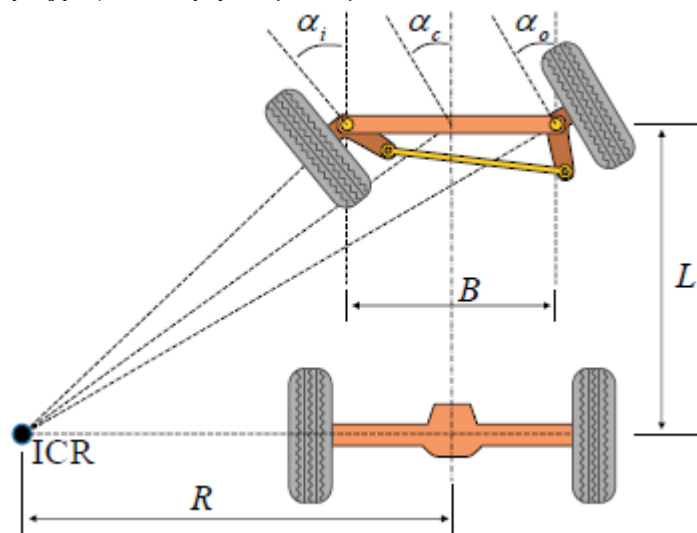


Рис. 5. Принцип рулевого управления Аккермана, позволяющий минимизировать трение при поворотах четырехколесного транспортного средства. ICR - текущий центр поворота, R - радиус поворота, B - расстояние между боковыми колесами, L - расстояние между передними и задними колесами, α_i и α_o - углы передних колес к текущему центру поворота.

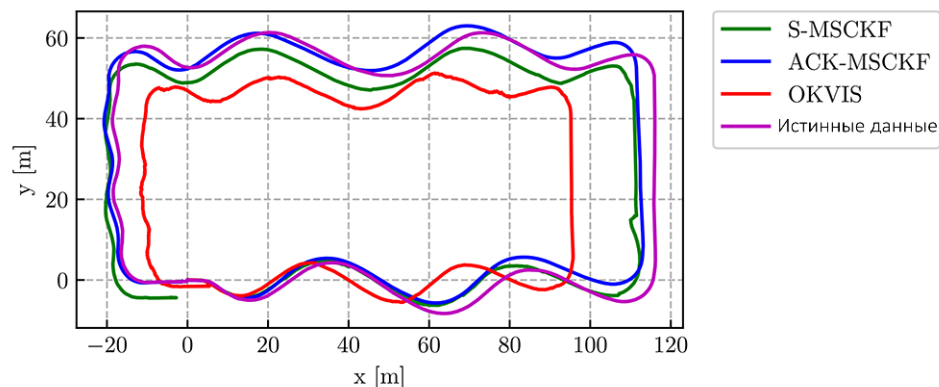


Рис. 6. Результат работы алгоритма ACK-MSCKF [] с принципом рулевого управления Аккермана и его сравнения с другими алгоритмами

6. Завершение

В данной работе проведен обзор современных методов в области визуально-инерциальной одометрии. Были рассмотрены различные подходы и тенденции развития систем визуально-инерциальной одометрии. Представлены схемы и формулы стандартного подхода к проблеме.

Литература

1. RTAB-Map as an Open-Source Lidar and Visual SLAM Library for Large-Scale and Long-Term Online Operation / M. Labbé, F. Michaud // *Journal of Field Robotics* – 2019 – vol. 36, no. 2 – p. 416–446.
2. DPPTAM: Dense piecewise planar tracking and mapping from a monocular sequence / A. Concha, J. Civera // In *IEEE International Conference on Intelligent Robots and Systems (IROS)* – 2015
3. LSD-SLAM: Large-Scale Direct Monocular SLAM / J. Engel, T. Schops, D. Cremers // In *IEEE Conference on European Conference on Computer Vision (ECCV)* – 2014 – p. 834–849.
4. PL-SVO: Semi-direct monocular visual odometry by combining points and line segments / R. Gomez-Ojeda, J. Briales, J. Gonzalez-Jimenez // In *IEEE Conference on Intelligent Robots and Systems (IROS)* – 2016
5. NID-SLAM: Robust Monocular SLAM using Normalised Information Distance / G. Pascoe, W. Maddern, M. Tanner, P. Pinies, P. Newman // In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* – 2017
6. CNN-SLAM: Real-Time Dense Monocular SLAM with Learned Depth Prediction / K. Tateno, F. Tombari, I. Laina, N. Navab // In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* – 2017
7. Robust Dense Mapping for Large-Scale Dynamic Environments / I. A. Barsan, P. Liu, M. Pollefeys, A. Geiger. // In *IEEE International Conference on Robotics and Automation (ICRA)* – 2018
8. Receding horizon «next-best-view» planner for 3d exploration / A. Bircher, M. Kamel, K. Alexis, H. Oleynikova, R. Siegwart // In *IEEE International Conference on Robotics and Automation (ICRA)* – 2016
9. StructVIO: Visual-inertial Odometry with Structural Regularity of Man-made Environments / D. Zou, Y. Wu, L. Pei, H. Ling, W. Yu // *IEEE Transactions on Robotics*, Vol. 35, Issue 4, 2019
10. Reconstruction of 3D Urban Scenes Using a Moving Lidar Sensor / O. Jozsa, C. Benedek // *Technical Report*, 2013
11. An Assessment of the Navigation and Course Corrections for a Manned Flyby of Mars or Venus / B. A. McElhoe // *IEEE Transactions on Aerospace and Electronic Systems*. 2 (4): 613–623, 1966
12. An efficient orientation filter for inertial and inertial/magnetic sensor arrays / S. O. H. Madjwick // *Internal report*, 2010
13. Applications of State-space Methods to Navigation Problems / S. Schmidt // In *Advances in Control Systems*. 3. New York, NY: Academic Press. pp. 293–340. 1966
14. A Review of Visual-Inertial Simultaneous Localization and Mapping from Filtering-Based and Optimization-Based Perspectives / C. Chen, H. Zhu, M. Li, S. You // In *Robotics* 2018, 7, 45

15. Visual-Inertial-Aided Navigation for High-Dynamic Motion in Built Environments without Initial Conditions / T. Lupton, S. Sukkarieh // In IEEE Transactions on Robotics, vol. 28, no. 1, pp. 61 – 76, Feb. 2012
16. A Multi-State Constraint Kalman Filter for Vision-aided Inertial Navigation / A. I. Mourikis, S. I. Roumeliotis // In IEEE International Conference on Robotics and Automation (ICRA), Roma, Italy, 10–14 April 2007; pp. 3565–3572
17. Visual-Inertial Navigation: A Concise Review / G. P. Huang // In IEEE International Conference on Robotics and Automation (ICRA) – 2019
18. Maplab: An Open Framework for Research in Visual-inertial Mapping and Localization / T. Schneider, M. Dymczyk, M. Fehr, K. Egger, S. Lynen, I. Gilitschenski, R. Siegwart // In IEEE International Conference on Robotics and Automation (ICRA) – 2018
19. Robust Stereo Visual Inertial Odometry for Fast Autonomous Flight / Ke Sun, K. Mohta, B. Pfrommer, M. Watterson // In IEEE Robotics and Automation Letters, 2018
20. ACK-MSCKF: Tightly-Coupled Ackermann Multi-State Constraint Kalman Filter for Autonomous Vehicle Localization / F. Ma, J. Shi, Y. Yang, J. Li, K. Dai // Sensors (Basel). 2019. Nov 5, 19 (21)
21. VINS-Mono: A Robust and Versatile Monocular Visual-Inertial State Estimator / T. Qin, P. Li, Z. Yang, S. Shen // In IEEE Transactions on Robotics, 2017
22. Keyframe-based visual-inertial odometry using nonlinear optimization / S. Leutenegger, S. Lynen, M. Bosse, R. Siegwart, P. T. Furgale // The International Journal of Robotics Research, 2015
23. Robocentric Visual-Inertial Odometry / Z. Huai, G. Huang // In IEEE Conference on Intelligent Robots and Systems (IROS) – 2018