

В.А. Сойфер, Е.В. Гошин

**ВИЗУАЛЬНАЯ ОДОМЕТРИЯ С КОРРЕКЦИЕЙ ТРАЕКТОРИИ
НА ОСНОВЕ ПРИМЕНЕНИЯ ФИЛЬТРА КАЛМАНА**

Сойфер В.А., Гошин Е.В. Визуальная одометрия с коррекцией траектории на основе применения фильтра Калмана.

Аннотация. В настоящей работе предложен метод коррекции траектории камеры в задаче монокулярной визуальной одометрии на основе фильтра Калмана. Рассмотрен подход, в котором неоднозначность масштаба разрешается не через восстановление трёхмерной структуры сцены, а через последовательное использование относительных смещений между текущим и несколькими предыдущими положениями камеры. Такой подход снижает зависимость от долговременного отслеживания одних и тех же ключевых точек и делает метод применимым в условиях ограниченного параллакса, повторяющихся текстур и частичных оклюзий. Подобное представление повышает чувствительность метода к шумам. Для компенсации этого недостатка применяется фильтр Калмана, который предотвращает накопление ошибок при последовательных оценках и обеспечивает корректность работы даже при отсутствии сведений о движении камеры, а также позволяет учитывать качество измерений и их достоверность. В рамках работы представлены модель состояния и модель наблюдений, позволяющие уточнять положение камеры, используя только данные последовательных наблюдений. Состояние формируется в виде скользящего окна из текущего и нескольких предыдущих положений камеры, а наблюдения строятся из направлений относительных смещений, оцениваемых по паре или набору кадров, без восстановления трёхмерной структуры сцены. В отличие от распространённых фильтров Калмана с ограничениями на множественные состояния системы (MSCKF), которые обычно интегрируют данные от инерциальных датчиков, предложенный метод использует исключительно информацию, извлекаемую из последовательности кадров. Эффективность метода подтверждена на синтетических данных в различных сценариях движения.

Ключевые слова: визуальная одометрия, фильтр Калмана, монокулярное зрение, оценка положения, робототехника.

1. Введение. Визуальная одометрия [1 – 3] представляет собой задачу оценки собственного движения камеры или робота на основе последовательности изображений. Точность таких оценок критически важна для автономной навигации [4 – 6], включая управление беспилотными летательными аппаратами [7, 8] и мобильными платформами. Традиционные методы, такие как восстановление структуры из движения (Structure from Motion, SfM) [9] и одновременная локализация и построение карты (Simultaneous Localization and Mapping, SLAM) [10], активно исследуются в последние десятилетия [11]. Известные методы оптимизации траектории, такие как групповая корректировка (bundle adjustment) [12], позволяют минимизировать ошибки репроекции, однако высокая вычислительная сложность ограничивает их применение в реальном времени [13].

Визуальная одометрия исторически развивалась в русле двух парадигм – прямых и признаковых методов – с последующим появлением гибридных подходов. Прямые методы минимизируют фотометрическую ошибку между кадрами, что обеспечивает высокую точность при стабильных условиях освещения и достаточном параллаксе. Признаковые методы опираются на детекторы и дескрипторы (например, SIFT/ORB/AKAZE), что делает их устойчивыми к умеренным фотометрическим и геометрическим искажениям, но чувствительными к деградации текстуры. Гибридные методы стремятся объединить преимущества обеих групп, дополняя геометрию фотометрией и наоборот, что особенно важно в динамических сценах и при малом параллаксе.

Задачи SfM/SLAM и одометрии находятся в тесной связи, но отличаются постановкой: SLAM нацелен на совместную оценку траектории и карты, тогда как одометрия обычно ограничивается локальной последовательной оценкой положения и ориентации камеры без явного построения глобальной карты. Это различие приводит к компромиссу между точностью и вычислительной сложностью: в то время как глобальные оптимизации (bundle adjustment) обеспечивают минимизацию ошибок репроекции на больших окнах, они труднее масштабируются для строго реального времени на ограниченных вычислительных ресурсах. В этой связи модели со скользящим окном и рекурсивной фильтрацией выглядят практически привлекательными.

В связи с этим активно исследуются подходы, позволяющие повысить точность и устойчивость при меньших вычислительных затратах. Одним из таких подходов является использование фильтра Калмана [14 – 16], который обеспечивает рекурсивную оценку состояния системы с учётом неопределённостей процесса и измерений. Как правило, фильтр Калмана применяется в составе визуально-инерциальных систем [17, 18], а также в сочетании с данными лидара [19], что обеспечивает высокую точность и устойчивость, но делает систему более сложной. Наибольшее распространение получили фильтры Калмана с ограничениями на множественные состояния системы (multi-state constraint Kalman filter, MSCKF) [20], где объединяются данные от нескольких последовательных положений камеры и инерциальных датчиков. Эти методы демонстрируют высокую точность, но требуют наличия и калибровки инерциальных датчиков, что повышает требования к аппаратной части.

С практической точки зрения, зависимость от инерциальных датчиков приводит к дополнительным этапам межсенсорной

калибровки, синхронизации и подавления смещений и дрейфа акселерометров и гироскопов. Более того, для лёгких и энергоограниченных платформ (микро-БПЛА, наземные мобильные микророботы) минимизация числа сенсоров часто является ключевым требованием. Поэтому варианты, использующие исключительно визуальные данные при сохранении устойчивости оценок, представляют особую ценность для реальных применений.

В настоящей работе рассматривается метод, использующий только визуальную информацию. Он сохраняет идею MSCKF – учитывать несколько последовательных положений камеры, – но устраняет зависимость от дополнительных сенсоров. Такой подход позволяет:

- работать исключительно с изображениями;
- адаптивно учитывать качество измерений (например, число верно сопоставленных точек после применения метода RANSAC [21], ошибку репроекции, величину параллакса);
- обеспечивать устойчивость даже в условиях сложных сцен, когда отслеживание одних и тех же ключевых точек на последовательных кадрах затруднено.

Работа демонстрирует применение фильтра Калмана для задачи коррекции траектории, полученной в результате визуальной одометрии. Предлагается модель состояния в виде скользящего окна из текущего и нескольких предыдущих положений камеры, что позволяет учитывать краткосрочную динамику без явного ввода скоростей и ускорений. Формируется модель наблюдений, основанная на согласовании направлений относительных смещений, оцениваемых по набору кадров, что устраняет необходимость восстановления трёхмерной структуры сцены и уменьшает зависимость от долговременного отслеживания одних и тех же точек. Вводится схема адаптивного взвешивания наблюдений по метрикам качества сопоставления, позволяющая повышать робастность к выбросам и шумам в визуальных данных. Метод ориентирован на работу в реальном времени и не требует инерциальных сенсоров, что снижает аппаратную сложность системы.

Помимо методологической значимости, подход имеет практическую ценность для задач автономной навигации в условиях отсутствия GPS/GNSS (склады, шахты, помещения со сложной геометрией), а также при ограниченном параллаксе (движение вдоль оси оптического луча, повторяющиеся или однородные текстуры). В таких сценариях отказ от восстановления плотной карты и использование направлений смещений в сочетании с фильтрацией по Калману создают устойчивую альтернативу тяжёлым глобальным

оптимизациям. В настоящей работе метод демонстрируется на задаче уточнения положения; при этом общая постановка распространяется и на оценку ориентации камеры.

2. Описание задачи. Целью исследования является разработка метода коррекции траектории камеры на основе последовательности монокулярных изображений с использованием фильтра Калмана. Входными данными служат оценки относительного движения, полученные стандартными методами определения положения и ориентации камеры по паре кадров (например, через фундаментальную или существенную матрицу). Эти оценки, как правило, надёжно задают направление смещения между положениями камеры, но не определяют масштаб (величину смещения).

Предполагается, что камера жёстко закреплена на движущемся носителе (рисунок 1). Для корректной работы монокулярной визуальной одометрии признаки желательны к наблюдению не менее чем в трёх последовательных кадрах. На практике это требование регулярно нарушается (динамические сцены, слабоструктурированные поверхности, изменения освещения), из-за чего традиционная схема [3], предполагающая восстановление трёхмерных точек, быстро теряет устойчивость.

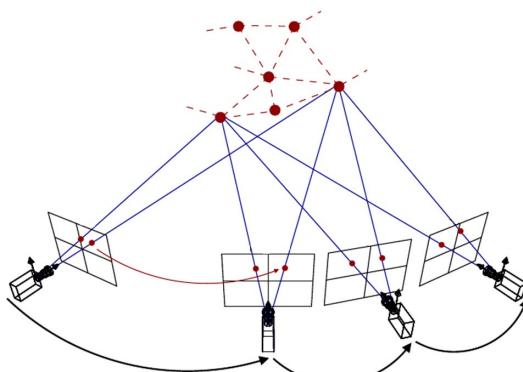


Рис. 1. Модель съёмки

Как правило, процедура оценки траектории движения состоит из следующих шагов:

1. Извлечение признаков на первом и втором кадрах.
2. Сопоставление признаков между кадрами и оценка относительного поворота и смещения (например, 5-точечным алгоритмом). Так как смещение задано с точностью до масштаба, на

в этом этапе вводится произвольный масштаб между первыми двумя положениями.

3. Извлечение признаков на следующем (третьем) кадре и сопоставление с ранее отобранными признаками.

4. Оценка относительного движения между текущим положением и несколькими предыдущими (обычно двумя и более), получая направления смещений из каждого предыдущего положения к текущему (рисунок 2).

5. Коррекция текущих координат камеры по информации, накопленной к текущему моменту.

6. Переход к следующему кадру и повторение процедуры, начиная с шага 4.

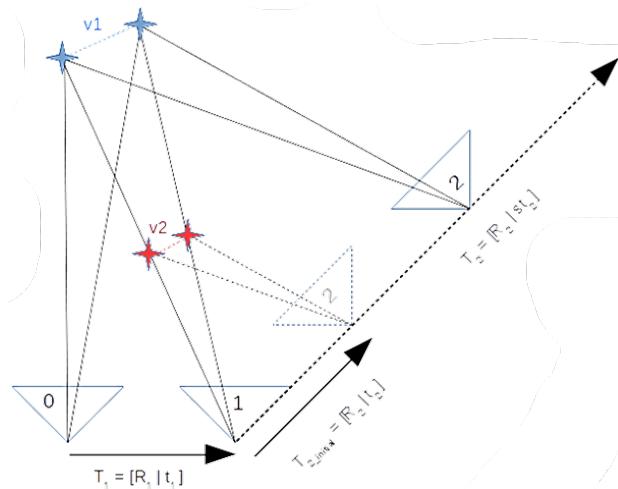


Рис. 2. Коррекция положения камеры

Отличие настоящей работы заключается в том, что на шаге 5 производится согласование направлений из нескольких предыдущих положений. Масштаб между первыми двумя кадрами задаётся произвольно. Далее для каждого последующего кадра рассматриваются несколько направлений смещения, оценённых относительно нескольких предыдущих положений (например, из положения на двух предыдущих к текущему). Геометрически это соответствует согласованию нескольких лучей: искомое положение для текущего положения выбирается как точка, наиболее согласующаяся с лучами, выходящими из предыдущих положений вдоль оценённых направлений. Такой подход не требует

восстановления трёхмерных точек сцены и не опирается на сохранение одних и тех же ключевых точек в соседних кадрах, что делает его применимым в трудных для отслеживания сценах (низкая текстура, повторяющиеся паттерны, быстро меняющиеся детали).

Как правило, в литературе этот способ редко описывается и применяется, поскольку чувствителен к угловым ошибкам направлений и к флуктуациям оценок относительных смещений. В настоящей работе этот недостаток компенсируется фильтром Калмана, который сглаживает случайные отклонения и препятствует накоплению ошибок в траектории, работает исключительно по визуальным данным, не требуя наличия инерциальных датчиков или других дополнительных сенсоров, а также учёт качества измерений, например, по доле выбросов после RANSAC.

3. Описание метода

Уравнение состояния. Для описания движения камеры будем использовать вектор состояния, включающий несколько предыдущих положений. Пусть положение камеры в момент времени t задаётся вектором $\mathbf{p}_t \in \mathbb{R}^3$.

Тогда вектор состояния определяется как

$$\mathbf{x}_t = \begin{bmatrix} \mathbf{p}_t \\ \mathbf{p}_{t-1} \\ \mathbf{p}_{t-2} \\ \vdots \\ \mathbf{p}_{t-k} \end{bmatrix} = \begin{bmatrix} \mathbf{X}_t \\ \mathbf{Y}_t \\ \mathbf{Z}_t \\ \mathbf{X}_{t-1} \\ \mathbf{Y}_{t-1} \\ \mathbf{Z}_{t-1} \\ \vdots \\ \mathbf{X}_{t-k} \\ \mathbf{Y}_{t-k} \\ \mathbf{Z}_{t-k} \end{bmatrix} \in \mathbb{R}^{3(k+1)}.$$

В отличие от традиционных схем, в нашем случае скорость и ускорение не включаются в явном виде в состояние системы, а при необходимости извлекается из последовательности положений камеры. Так, например, скорость может быть извлечена через конечные разности как:

$$\hat{\mathbf{v}}_t = \mathbf{p}_t - \mathbf{p}_{t-1}.$$

Соответственно, прогноз текущего положения выполняется по формуле линейной экстраполяции

$$\mathbf{p}_{t+1} \approx \mathbf{p}_t + \hat{\mathbf{v}}_t = 2\mathbf{p}_t - \mathbf{p}_{t-1}.$$

Уравнение состояния описывает эволюцию системы во времени и записывается в виде:

$$\mathbf{x}_t = \mathbf{F}\mathbf{x}_{t-1} + \mathbf{w}_t,$$

где \mathbf{F} – матрица перехода, экстраполирующая новое состояние по предыдущим данным, \mathbf{w}_k – процессный шум, моделирующий неопределённость прогноза (например, из-за изменения скорости или манёвров).

Поскольку состояние \mathbf{x}_t состоит из текущей позиции и k предыдущих, структура \mathbf{F} естественным образом блочная. Очевидно, что для нового состояния \mathbf{x}_{t+1} положения камеры в предыдущие моменты времени $\mathbf{p}_{t+1-i}, i = 1, k$ переносятся из предыдущего шага без изменений, то есть нижняя часть матрицы содержит единичные 3×3 блоки на поддиагонали, обеспечивая сдвиг состояний.

Текущее положение экстраполируется с использованием двух последних точек траектории $\mathbf{p}_t \approx 2\mathbf{p}_{t-1} - \mathbf{p}_{t-2}$. Соответственно, верхняя блок-строка \mathbf{F} имеет ненулевые коэффициенты при блоках, соответствующих \mathbf{p}_{t-1} и \mathbf{p}_{t-2} (например, $2\mathbf{I}$ и $-\mathbf{I}$) тогда как остальные блоки в этой строке равны нулю. Таким образом, матрица \mathbf{F} приобретает вид

$$\mathbf{F} = \begin{bmatrix} 2\mathbf{I} & -\mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{I} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I} & \mathbf{0} \end{bmatrix},$$

где \mathbf{I} – единичная матрица размерности 3×3 , $\mathbf{0}$ – нулевая матрица размерности 3×3 .

Такое построение матрицы перехода одновременно остаётся простым вычислительно и позволяет учитывать краткосрочную динамику движения без явного введения скоростей и ускорений в состояние. За счёт процессного шума \mathbf{w} модель остаётся гибкой и устойчивой к отклонениям от идеальной модели постоянной скорости.

При необходимости точность прогноза можно повысить введением дополнительных слагаемых, учитывающих ускорение или повороты камеры. Следует добавить, что предложенная схема

позволяет гибко увеличивать окно учитываемых позиций в зависимости от доступной памяти и вычислительных возможностей. Например, при наличии аппаратного ускорения (параллельной обработки или GPU) можно расширить окно до 6–8 позиций, что позволяет более надёжно сглаживать шум и предсказывать траекторию в условиях низкого параллакса. На практике размер окна следует выбирать как компромисс между точностью и допустимой задержкой в конкретной системе реального времени.

Ковариационная матрица процесса. Ковариационная матрица ошибки состояния \mathbf{P}_t отражает нашу уверенность в прогнозе и на каждом шаге пересчитывается с учётом динамики модели. В рамках фильтра Калмана это делается по стандартной формуле

$$\mathbf{P}_t = \mathbf{F}\mathbf{P}_{t-1}\mathbf{F}^T + \mathbf{Q}_t,$$

где первый член переносит накопленную неопределённость через матрицу перехода \mathbf{F} , а \mathbf{Q}_t моделирует неточность прогноза, связанную с неожиданными изменениями движения.

Учитывая структуру состояния (текущее и k предыдущих положений), естественно задавать \mathbf{Q}_t так, чтобы шум процесса влиял прежде всего на текущую компоненту, а прошлые позиции лишь переносились со своей уже накопленной неопределённостью. Практически это выражается в блочно-диагональной форме

$$\mathbf{Q}_t = \begin{bmatrix} q\mathbf{I} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

где \mathbf{I} – единичная матрица размерности 3×3 , а $q > 0$ – коэффициент, отвечающий за погрешность прогнозирования.

Чем более изменчивым ожидается движение (резкие ускорения, манёвры), тем большим следует выбирать q ; при плавной, близкой к равномерной траектории q можно уменьшать. Так, если движение камеры предсказуемо и плавно (например, при съёмке с робота на колёсах), величина дополнительного шумового члена может быть уменьшена для повышения точности оценки. В противоположность этому, при резких манёврах разумно увеличить ковариацию процесса, чтобы фильтр быстрее реагировал на новые измерения.

Такой выбор \mathbf{Q}_t позволяет явно учитывать неопределенность в экстраполяции текущего положения, не увеличивая ковариацию для уже наблюдённых поз прошлого окна, а также сохранять простоту модели и стабильность оценки, поскольку дополнительные корреляции между блоками вносит сама операция $\mathbf{F}\mathbf{P}_{t-1}\mathbf{F}^T$.

При наличии априорных сведений о направлении большей вариабельности (например, ограничения по высоте или доминирующее движение в плоскости) матрицу \mathbf{Q}_t можно сделать анизотропной, заменив $\mathbf{Q}_t = q\mathbf{I}$ на $\mathbf{Q}_t = \text{diag}(q_x, q_y, q_z)$ с различными дисперсиями по осям. Если шум естественное задавать в системе камеры, допустимо учитывать ориентацию и задавать $\mathbf{Q}_t = \mathbf{R}\text{diag}(q_x, q_y, q_z)\mathbf{R}^T$ для текущего блока. В любом случае ключевая идея сохраняется: на каждом шаге ковариация процесса добавляется прежде всего к компоненте, отвечающей за текущее положение, что позволяет фильтру гибко реагировать на изменения характера движения.

Уравнение наблюдения. Уравнение наблюдения связывает состояние системы с измерениями:

$$\mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{v}_t,$$

где:

\mathbf{x}_t – вектор измерений (проекции точек сцены на изображение);

\mathbf{H}_t – матрица наблюдения;

\mathbf{v}_t – шум измерений.

Пусть в каждый новый момент времени t методами визуальной одометрии выполняется оценка относительного смещения текущего положения камеры относительно нескольких предыдущих. На практике визуальная одометрия надёжно определяет направление смещения \mathbf{d} (как правило, нормированное, $\|\mathbf{d}\| = 1$) от положения $\mathbf{p}_i, i \in \{t-1, t-k\}$ к новому положению \mathbf{p}_i , но абсолютное значение (масштаб) этого смещения неизвестно.

Таким образом, задана прямая, на которой расположены новые координаты, через предыдущие известные координаты \mathbf{p}_{t-i} и направление $\mathbf{d}_{t,i}$ с точностью до неизвестного коэффициента α :

$$\mathbf{p}_t = \mathbf{p}_{t-i} + \alpha \mathbf{d}_{t,i},$$

Рассмотрим проекцию

$$\text{proj}_{\mathbf{d}_{t,i}}(\mathbf{p}_t - \mathbf{p}_{t-i}) = \frac{(\mathbf{p}_t - \mathbf{p}_{t-i}) \cdot \mathbf{d}_{t,i}}{\mathbf{d}_{t,i} \cdot \mathbf{d}_{t,i}} \mathbf{d}_{t,i}$$

вектора смещения $\mathbf{p}_t - \mathbf{p}_{t-i}$ на направление $\mathbf{d}_{t,i}$. В случае, если направление $\mathbf{d}_{t,i}$ и векторы смещения вычислены точно, эта проекция равна:

$$\text{proj}_{\mathbf{d}_{t,i}}(\mathbf{p}_t - \mathbf{p}_{t-i}) = \frac{(\alpha \mathbf{d}_{t,i}) \cdot \mathbf{d}_{t,i}}{\mathbf{d}_{t,i} \cdot \mathbf{d}_{t,i}} \mathbf{d}_{t,i} = \alpha \mathbf{d}_{t,i} = \mathbf{p}_t - \mathbf{p}_{t-i}.$$

Таким образом, для каждой пары $\mathbf{p}_t, \mathbf{p}_{t-i}$ можно задать наблюдение в момент времени t в виде:

$$\mathbf{h}_{t,i} = \text{proj}_{\mathbf{d}_{t,i}}(\mathbf{p}_t - \mathbf{p}_{t-i}) - (\mathbf{p}_t - \mathbf{p}_{t-i}) = (\mathbf{p}_t - \mathbf{p}_{t-i}) \cdot \mathbf{d}_{t,i} \cdot \mathbf{d}_{t,i} - (\mathbf{p}_t - \mathbf{p}_{t-i}) = (\mathbf{d}_{t,1} \mathbf{d}_{t,1}^T - \mathbf{I}) \mathbf{p}_t - (\mathbf{d}_{t,1} \mathbf{d}_{t,1}^T - \mathbf{I}) \mathbf{p}_{t-i} = \mathbf{0}.$$

Для каждого предыдущего состояния формируется одно наблюдение такого рода, в результате чего итоговая матрица наблюдения приобретает вид:

$$\mathbf{H}_k = \begin{bmatrix} \mathbf{d}_{t,1} \mathbf{d}_{t,1}^T - \mathbf{I} & -\mathbf{d}_{k-1} \mathbf{d}_{k-1}^T + \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{d}_{t,2} \mathbf{d}_{t,2}^T - \mathbf{I} & \mathbf{0} & -\mathbf{d}_{k-2} \mathbf{d}_{k-2}^T + \mathbf{I} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{d}_{t,k} \mathbf{d}_{t,k}^T - \mathbf{I} & \mathbf{0} & \mathbf{0} & \cdots & -\mathbf{d}_{t,k} \mathbf{d}_{t,k}^T + \mathbf{I} \end{bmatrix}.$$

Такая модель наблюдений удобна для интеграции с существующими блоками визуальной одометрии, поскольку не требует модификации внутренней логики сопоставления признаков. Фактически она позволяет использовать уже имеющиеся результаты относительных оценок движения и лишь дополнять их статистическим уточнением через фильтр Калмана.

Ковариационная матрица шума измерений. Поскольку наблюдения формируются на основе результатов визуальной одометрии и неизбежно содержат ошибки (неточности сопоставления, угловые ошибки направлений, неопределенность масштаба), их достоверность необходимо учесть. Ковариационная матрица шума измерений \mathbf{V}_t задается как

$$\mathbf{R}_t = \begin{bmatrix} \sigma_{t-1}\mathbf{I} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sigma_{t-2}\mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sigma_{t-3}\mathbf{I} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \sigma_{t-k}\mathbf{I} \end{bmatrix},$$

где σ_{t-i}^2 – дисперсия i -го скалярного измерения z_i . Тем самым каждое наблюдение получает собственный вес в зависимости от качества оценки.

Практически σ_{t-i}^2 могут быть одинаковыми при отсутствии информации о качестве визуальной одометрии, либо выбираться на основе доступных метрик, например:

- количество сопоставленных ключевых точек и доля выбросов (после RANSAC);
- средняя или медианная ошибка репроекции;
- оценка неопределенности существенной или фундаментальной матрицы;
- величина параллакса (больший параллакс – меньшая неопределенность).

Нормировка направлений $\mathbf{d}_{t,i}$ делает наблюдения z_i сопоставимыми по шкале, что упрощает калибровку \mathbf{R}_t . При наличии априорной информации допускается адаптивная схема: \mathbf{R}_t пересчитывается на каждом шаге по выбранным метрикам качества, благодаря чему фильтр автоматически уменьшает влияние ненадежных наблюдений и усиливает вклад более точных. Такой механизм повышает устойчивость оценки в условиях переменного уровня шума и наличия выбросов.

В итоге совокупность представленных уравнений состояния, наблюдений и ковариаций образует полную схему фильтра Калмана, предназначенного для визуальной одометрии без инерциальных данных. Такое построение обеспечивает баланс между простотой реализации и достаточной точностью, позволяя применять метод как на мощных вычислительных платформах, так и на встраиваемых устройствах.

4. Экспериментальное исследование на синтетических данных. Для демонстрации принципиальной работоспособности подхода использовались генерированные данные. Такой формат позволяет контролировать траекторию движения, уровень шума и параметры наблюдений, а значит – корректно сравнивать методы в одинаковых условиях.

Моделирование движения аппарата выполнялось в трёх сценариях:

- линейная траектория с незначительными отклонениями;
- круговая траектория;
- произвольная траектория со случайным изменением скорости и направления.

Для каждой из этих траекторий моделирование движения аппарата производилось следующим образом.

Линейная траектория с незначительными отклонениями:

$$\mathbf{p}_{i+1} = \mathbf{p}_i + \Delta\mathbf{p} \cdot v_i + \mathcal{N}(0, \sigma_p^2),$$

$$v_{i+1} = v_i + \mathcal{N}(0, \sigma_v^2),$$

где \mathbf{p}_i – точка траектории, $\Delta\mathbf{p}$ – направление движения, v_i – абсолютное значение скорости камеры, σ_p^2 – дисперсия отклонения от траектории, σ_v^2 – дисперсия скорости. Изначальная скорость v_i задаётся при моделировании. Начальное положение точки задано как $(0, 0, 0)$. Такая траектория позволяет оценить качество траектории для случаев, когда соответствующие лучи практически коллинеарны.

Круговая или эллиптическая траектория:

$$\mathbf{p}_i = c + (\cos(\psi_i) \cdot R_x + \sin(\psi_i) \cdot R_y) + \mathcal{N}(0, \sigma_p^2),$$

$$\psi_i = \psi_{i+1} + v_i,$$

$$v_{i+1} = v_i + \mathcal{N}(0, \sigma_v^2),$$

где \mathbf{p}_i – точка траектории, c – центр круга или эллипса, R_x, R_y – полуоси эллипса, ψ_i – угловое положение камеры, v_i – угловая скорость камеры, σ_p^2 – дисперсия отклонения от траектории, σ_v^2 – дисперсия скорости. Изначальная скорость v_i задаётся при моделировании.

Произвольная траектория со случайным изменением направления и абсолютного значения скорости:

$$\mathbf{p}_{i+1} = \mathbf{p}_i + \mathbf{v}_i + \mathcal{N}(0, \sigma_p^2),$$

$$\mathbf{v}_{i+1} = \mathbf{v}_i + \mathcal{N}(0, \sigma_v^2),$$

где \mathbf{p}_i – точка траектории, \mathbf{v}_i – скорость камеры, σ_p^2 – дисперсия отклонения от траектории, σ_v^2 – дисперсия скорости. Изначальная скорость v_i задаётся при моделировании. Начальное положение точки

задано как $(0, 0, 0)$. Такая траектория позволяет оценить качество траектории для случаев, когда скорость значительно меняется во времени как по амплитуде, так и по направлению, вследствие чего погрешность прогноза нового состояния системы велика.

С использованием известных положений камеры были вычислены точные значения направлений $\mathbf{d}_{t,i}$, после чего они зашумлялись аддитивным гауссовым шумом, моделирующим погрешности визуальной одометрии (ошибки сопоставления). Интенсивность шума варьировалась через параметр SNR.

Было рассмотрено два подхода к определению следующего положения камеры:

1. Метод минимизации расстояний. Используется к предыдущих положений аппарата и направление смещения относительно каждого. Минимизируется сумма расстояний от соответствующих прямых.

$$\mathbf{p}_t = \min \left\| \sum dist(l_{t,k}, \mathbf{p}) \right\|,$$

$l_{t,k}$ – прямая, заданная точкой \mathbf{p}_{t-k} и направлением $\mathbf{d}_{t,k}$, $dist(l_{t,k}, \mathbf{p})$ – расстояние от точки до прямой.

2. Предложенный метод (фильтр Калмана). Используется к предыдущих положений аппарата и направление смещение относительно каждого. Применяется фильтр Калмана для уточнения положения.

Также иногда применяется проход, при котором используется только одно предыдущее положение аппарата и направление смещения. Поскольку при изменении скорости такой подход демонстрирует нестабильность и существенные ошибки, в ходе эксперимента он не использовался.

В качестве метрики качества далее используется отклонение по всем точкам траектории; анализ проводится для всех трёх типов движения и различных уровней шума. Далее будут представлены результаты сравнения традиционного и предложенного методов в условиях неизвестных свойств шума.

Оценка качества проводилась на синтетических траекториях для трёх сценариев движения: линейного с незначительными отклонениями, кругового и произвольного с изменяющейся скоростью и направлением. Для каждого сценария рассматривались различные уровни шума наблюдений ($SNR = 50, 60, 70$). В качестве метрики использовалось среднее отклонение по всем точкам траектории.

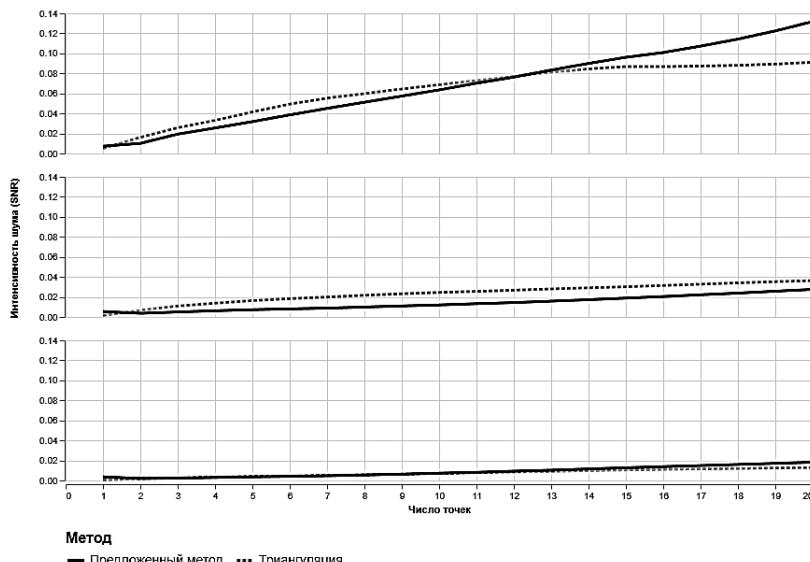


Рис. 3. Линейная траектория с незначительными отклонениями, зависимость среднего отклонения от номера кадра (сверху вниз, SNR: 50, 60, 70)

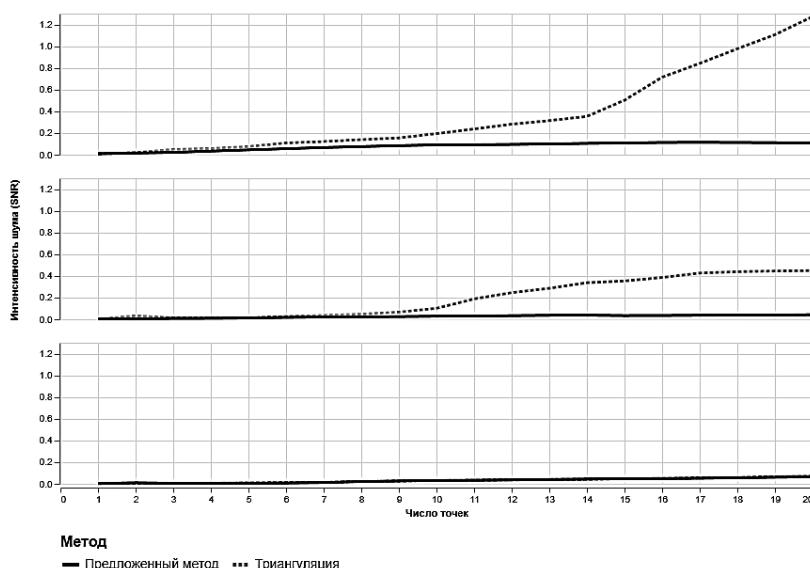


Рис. 4. Круговая траектория, зависимость среднего отклонения от номера кадра (сверху вниз, SNR: 50, 60, 70)

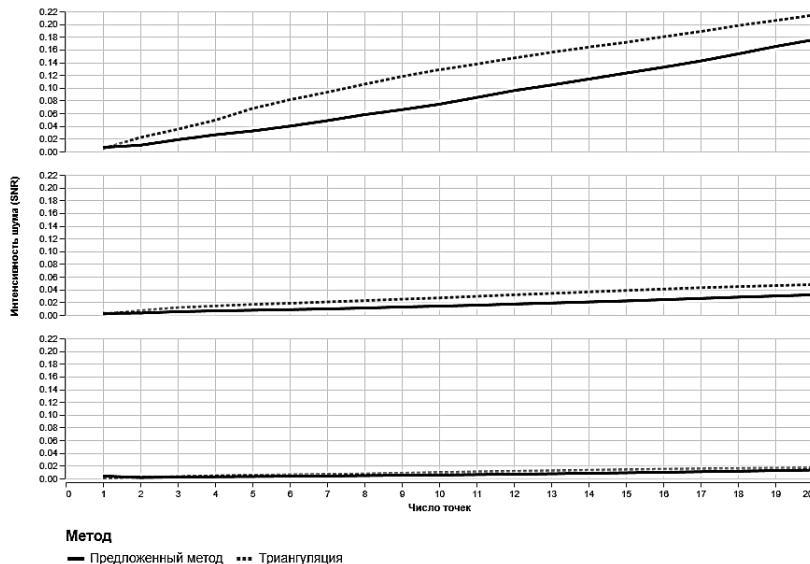


Рис. 5. Произвольная траектория со случайным изменением скорости, зависимость среднего отклонения от номера кадра (сверху вниз, SNR: 50, 60, 70)

Экспериментальное исследование показало, что предложенный метод на основе фильтра Калмана обеспечивает более устойчивую и точную оценку положения по сравнению с традиционной триангуляцией. Наиболее заметное преимущество наблюдается на круговой траектории и при высоких уровнях шума, что указывает на рабочность предлагаемого решения к неблагоприятным условиям.

5. Заключение. В данной работе предложен метод коррекции траектории камеры в задаче визуальной одометрии с использованием фильтра Калмана. Предложены модель состояний, матрица перехода и модель наблюдений, обеспечивающие решение поставленной задачи.

Экспериментально показано, что предложенный метод обеспечивает более высокую точность оценки положения по сравнению с базовыми подходами, особенно в условиях переменного шума и наличия выбросов в данных. Это делает его перспективным для применения в реальных системах визуальной одометрии, включая автономные транспортные средства, БПЛА и мобильных роботов.

Основными достоинствами предложенного метода являются:

- Возможность работы только с визуальными данными, без необходимости использования дополнительных сенсоров.

- Адаптивная настройка параметров фильтра на основе оценки достоверности измерений.
- Высокая устойчивость к шумам и выбросам в данных.
- Возможность работы в реальном времени благодаря низкой вычислительной сложности.

В дальнейшем планируется провести тестирование в реальных условиях для оценки применимости метода в практических сценариях.

Для авторов большая часть участвовать данной статьёй в журнале, посвящённом памяти выдающегося учёного, члена-корреспондента РАН Рафаэля Мидхатовича Юсупова.

Тематика статьи относится к научным интересам Р.М. Юсупова, который проводил исследования в области теории управления полётом летательных аппаратов, теории адаптивных систем.

С Самарой и Самарским университетом имени академика С.П. Королева Рафаэля Мидхатовича связывали многолетние творческие и дружеские отношения. Будучи волжанином, Р.М. Юсупов часто бывал в Самаре, принимал участие в научных конференциях, общался с профессорско-преподавательским составом, научными работниками и студентами.

Творческое наследие Р.М. Юсупова способствует развитию научных исследований, а обаяние его личности остаётся в нашей памяти.

Литература

1. Neyestani A., Picariello F., Ahmed I., Daponte P., De Vito L. From pixels to precision: A survey of monocular visual odometry in digital twin applications // Sensors. 2024. vol. 24. no. 4.
2. Yuan S., Zhang J., Lin Y., Yang L. Hybrid self-supervised monocular visual odometry system based on spatio-temporal features // Electronic Research Archive. 2024. vol. 32. no. 5. pp. 3543–3568.
3. Scaramuzza D., Fraundorfer F. Visual odometry [tutorial] // IEEE Robotics & Automation Magazine. 2011. vol. 18. no. 4. p. 80–92.
4. Ou Y., Cai Y., Sun Y., Qin T. Autonomous navigation by mobile robot with sensor fusion based on deep reinforcement learning // Sensors. 2024. vol. 24. no. 12.
5. Лебедев А.О., Васильев В.В., Паулиш А.Г. Алгоритм управления полетом БПЛА вдоль железной дороги с помощью технического зрения // Компьютерная оптика. 2025. Т. 49. № 2. С. 320–326. DOI: 10.18287/2412-6179-CO-1532.
6. Belkin I.V., Abramenco A.A., Bezuglyi V.D., Yudin D.A. Localization of mobile robot in prior 3D LiDAR maps using stereo image sequence // Computer Optics. 2024. vol. 48. no. 3. pp. 406–417.
7. Горбачев В.А., Калугин В.Ф. Разработка алгоритма многообъектного трекинга с необучаемыми признаками сопоставления объектов // Компьютерная оптика. 2023. Т. 47. № 6. С. 1002–1010. DOI: 10.18287/2412-6179-CO-1275.
8. Ляхов П.А., Оразаев А.Р. Анализ видеоданных беспилотного летательного аппарата на основе индекса структурного сходства // Компьютерная оптика. 2025. Т. 49. № 4. С. 624–633. DOI: 10.18287/2412-6179-CO-1569.

9. Jiang S., Jiang C., Jiang W. Efficient structure from motion for large-scale UAV images: A review and a comparison of SfM tools // ISPRS Journal of Photogrammetry and Remote Sensing. 2020. vol. 167. pp. 230–251.
10. Le V.H. Visual SLAM and Visual Odometry Based on RGB-D Images Using Deep Learning: A Survey // Journal of Robotics and Control (JRC). 2024. vol. 5. no. 4. pp. 1050–1079.
11. Yan F., Li Z., Zhou Z. Robust and efficient edge-based visual odometry // Computational Visual Media. 2022. vol. 8. no. 3. pp. 467–481.
12. Weber S., Demmel N., Chan T.C., Cremers D. Power bundle adjustment for large-scale 3D reconstruction // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2023. pp. 281–289.
13. Triggs B., McLauchlan P., Hartley R., Fitzgibbon A. Bundle adjustment – A modern synthesis // Vision Algorithms: Theory and Practice. Lecture Notes in Computer Science. 2000. vol. 1883. pp. 298–372. DOI: 10.1007/3-540-44480-7_21.
14. Сойфер В.А., Фурсов В.А., Харитонов С.И. Калмановская фильтрация одного класса изображений динамических объектов // Информатика и автоматизация. 2024. Т. 23. № 4. С. 953–968. DOI: 10.15622/ia.23.4.1.
15. Tavares Jr A.J.A., Oliveira N.M.F. A Novel Approach for Kalman Filter Tuning for Direct and Indirect Inertial Navigation System/Global Navigation Satellite System Integration // Sensors. 2024. vol. 24. no. 22. DOI: 10.3390/s24227331.
16. Yuan Y., Li F., Chen J., Wang Y., Liu K. An improved Kalman filter algorithm for tightly GNSS/INS integrated navigation system // Mathematical Biosciences and Engineering. 2024. vol. 21. pp. 963–983.
17. Wang Z., Pang B., Song Y., Yuan X., Xu Q., Li Y. Robust visual-inertial odometry based on a Kalman filter and factor graph // IEEE Transactions on Intelligent Transportation Systems. 2023. vol. 24. no. 7. pp. 7048–7060.
18. Nguyen K.D., Tran D.T., Nguyen D.T., Inoue K., Lee J.H., Nguyen A.Q. Learning Visual-Inertial Odometry with Robocentric Iterated Extended Kalman Filter // IEEE Access. 2024. vol. 12. pp. 109943–109956.
19. Lee D., Jung M., Yang W., Kim A. Lidar odometry survey: recent advancements and remaining challenges // Intelligent Service Robotics. 2024. vol. 17. no. 2. pp. 95–118.
20. Xue C., Huang Y., Zhao C., Li X., Mihaylova L., Li Y., Chambers J.A. A Gaussian-generalized-inverse-Gaussian joint-distribution-based adaptive MSCKF for visual-inertial odometry navigation // IEEE Transactions on Aerospace and Electronic Systems. 2022. vol. 59. no. 3. pp. 2307–2328.
21. Fischler M.A., Bolles R.C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography // Communications of the ACM. 1981. vol. 24. no. 6. pp. 381–395.

Сойфер Виктор Александрович — д-р техн. наук, профессор, академик РАН, лауреат Государственной премии Российской Федерации в области науки и техники, президент Самарский национальный исследовательский университет имени академика С.П. Королёва. Область научных интересов: компьютерная оптика, обработка изображений, распознавание образов. Число научных публикаций — 871. soifer@ssau.ru; Московское шоссе, 34, 443086, Самара, Россия; р.т.: +7(846)332-2604.

Гошин Егор Вячеславович — канд. техн. наук, доцент, заведующий кафедрой, кафедра киберфотоники, Самарский национальный исследовательский университет имени академика С.П. Королёва. Область научных интересов: обработка изображений. Число научных публикаций — 78. goshine@ssau.ru; Молодогвардейская улица, 151, 443001, Самара, Россия; р.т.: +7(846)332-2606.

Поддержка исследований. Работа выполнена в рамках государственного задания по теме FSSS-2023-0006.

V. SOIFER, Y. GOSHIN

VISUAL ODOMETRY WITH TRAJECTORY CORRECTION BASED ON KALMAN FILTER

Soifer V., Goshin Y. Visual Odometry with Trajectory Correction Based on Kalman Filter.

Abstract. This paper proposes a Kalman-filter-based method for camera trajectory correction in monocular visual odometry. The approach resolves the scale ambiguity not by reconstructing the 3D structure of the scene, but by sequentially using relative displacements between the current and several previous camera positions. This mechanism reduces reliance on long-term tracking of the same keypoints and makes the method applicable under limited parallax, repetitive textures, and partial occlusions. However, this formulation increases the method's sensitivity to noise. To compensate for this drawback, a Kalman filter is employed, which prevents error accumulation during consecutive estimates, ensures correct operation even in the absence of information about camera motion, and allows measurement quality and reliability to be taken into account. A state model and an observation model are proposed, which allow for refining the camera pose using only data from consecutive observations. The state is represented as a sliding window comprising the current and several preceding camera poses, while the observations are constructed from directions of relative displacements estimated from an image pair or a short set of frames, without reconstructing the three-dimensional structure of the scene. Unlike common Multi-State Constraint Kalman Filters (MSCKF), which typically integrate inertial measurements, the proposed method relies exclusively on information extracted from the image sequence. The effectiveness of the proposed method is confirmed with synthetic data in various motion scenarios.

Keywords: visual odometry, Kalman filter, monocular vision, position estimation, robotics.

References

1. Neyestani A., Picariello F., Ahmed I., Daponte P., De Vito L. From pixels to precision: A survey of monocular visual odometry in digital twin applications. Sensors. 2024. vol. 24. no. 4.
2. Yuan S., Zhang J., Lin Y., Yang L. Hybrid self-supervised monocular visual odometry system based on spatio-temporal features. Electronic Research Archive. 2024. vol. 32. no. 5. pp. 3543–3568.
3. Scaramuzza D., Fraundorfer F. Visual odometry [tutorial]. IEEE Robotics & Automation Magazine. 2011. vol. 18. no. 4. p. 80–92.
4. Ou Y., Cai Y., Sun Y., Qin T. Autonomous navigation by mobile robot with sensor fusion based on deep reinforcement learning. Sensors. 2024. vol. 24. no. 12.
5. Lebedev A.O., Vasil'ev V.V., Paulish A.G. [Algorithm for UAV flight controlling along a railway using technical vision]. Komp'yuternaja optika – Computer Optics. 2025. vol. 49. no. 2. pp. 320–326. DOI: 10.18287/2412-6179-CO-1532. (In Russ.).
6. Belkin I.V., Abramenko A.A., Bezuglyi V.D., Yudin D.A. Localization of mobile robot in prior 3D LiDAR maps using stereo image sequence. Computer Optics. 2024. vol. 48. no. 3. pp. 406–417.
7. Gorbachev V.A., Kalugin V.F. [Development of a multi-object tracking algorithm with untrained features of object matching]. Komp'yuternaja optika – Computer Optics. 2023. vol. 47. no. 6. pp. 1002–1010. DOI: 10.18287/2412-6179-CO-1275. (In Russ.).
8. Ljahov P.A., Orazaev A.R. [Analysis of video data from an unmanned aerial vehicle based on a structural similarity index]. Komp'yuternaja optika – Computer Optics. 2025. vol. 49. no. 4. pp. 624–633. DOI: 10.18287/2412-6179-CO-1569. (In Russ.).

9. Jiang S., Jiang C., Jiang W. Efficient structure from motion for large-scale UAV images: A review and a comparison of SfM tools. *ISPRS Journal of Photogrammetry and Remote Sensing*. 2020. vol. 167. pp. 230–251.
10. Le V.H. Visual SLAM and Visual Odometry Based on RGB-D Images Using Deep Learning: A Survey. *Journal of Robotics and Control (JRC)*. 2024. vol. 5. no. 4. pp. 1050–1079.
11. Yan F., Li Z., Zhou Z. Robust and efficient edge-based visual odometry. *Computational Visual Media*. 2022. vol. 8. no. 3. pp. 467–481.
12. Weber S., Demmel N., Chan T.C., Cremers D. Power bundle adjustment for large-scale 3D reconstruction. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2023. pp. 281–289.
13. Triggs B., McLauchlan P., Hartley R., Fitzgibbon A. Bundle adjustment – A modern synthesis. *Vision Algorithms: Theory and Practice*. Lecture Notes in Computer Science. 2000. vol. 1883. pp. 298–372. DOI: 10.1007/3-540-44480-7_21.
14. Soifer V.A., Fursov V.A., Kharitonov S.I. [Kalman filtering of one class of dynamic object images]. *Informatics and Automation*. 2024. vol. 23. no. 4. pp. 953–968. DOI: 10.15622/ia.23.4.1. (In Russ.).
15. Tavares Jr A.J.A., Oliveira N.M.F. A Novel Approach for Kalman Filter Tuning for Direct and Indirect Inertial Navigation System/Global Navigation Satellite System Integration. *Sensors*. 2024. vol. 24. no. 22. DOI: 10.3390/s24227331.
16. Yuan Y., Li F., Chen J., Wang Y., Liu K. An improved Kalman filter algorithm for tightly GNSS/INS integrated navigation system. *Mathematical Biosciences and Engineering*. 2024. vol. 21. pp. 963–983.
17. Wang Z., Pang B., Song Y., Yuan X., Xu Q., Li Y. Robust visual-inertial odometry based on a Kalman filter and factor graph. *IEEE Transactions on Intelligent Transportation Systems*. 2023. vol. 24. no. 7. pp. 7048–7060.
18. Nguyen K.D., Tran D.T., Nguyen D.T., Inoue K., Lee J.H., Nguyen A.Q. Learning Visual-Inertial Odometry with Robocentric Iterated Extended Kalman Filter. *IEEE Access*. 2024. vol. 12. pp. 109943–109956.
19. Lee D., Jung M., Yang W., Kim A. Lidar odometry survey: recent advancements and remaining challenges. *Intelligent Service Robotics*. 2024. vol. 17. no. 2. pp. 95–118.
20. Xue C., Huang Y., Zhao C., Li X., Mihaylova L., Li Y., Chambers J.A. A Gaussian-generalized-inverse-Gaussian joint-distribution-based adaptive MSCKF for visual-inertial odometry navigation. *IEEE Transactions on Aerospace and Electronic Systems*. 2022. vol. 59. no. 3. pp. 2307–2328.
21. Fischler M.A., Bolles R.C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*. 1981. vol. 24. no. 6. pp. 381–395.

Soifer Victor — Ph.D., Dr.Sci., Professor, Academician of the RAS, Recipient of the Russian Federation State Prize in science and technology, President, Samara National Research University. Research interests: computer optics, image processing, pattern recognition. The number of publications — 871. soifer@ssau.ru; 34, Moscow Hwy, 443086, Samara, Russia; office phone: +7(846)332-2604.

Goshin Yegor — Ph.D., Associate Professor, Head of the department, Department of cyberphotronics, Samara National Research University. Research interests: image processing. The number of publications — 78. goshine@ssau.ru; 151, Molodogvardeyskaya St., 443001, Samara, Russia; office phone: +7(846)332-2606.

Acknowledgements. The research was carried out within the state assignment theme FSSS-2023-0006.