



Визуально-инерциальная навигация БПЛА с адаптивной оценкой неопределённости нейросетевой локализации на основе эвиденциального обучения

П.А. Лазарева, ORCID: 0000-0002-1393-3091 <palazareva@kai.ru>
Н.В. Гурько, ORCID: 0009-0001-0217-0471 <nickolay73@hotmail.com>

Казанский национальный исследовательский технический
университет им. А.Н. Туполева – КАИ,
Россия, 420111, г. Казань, ул. К.Маркса, д. 10.

Аннотация. Рассматривается задача визуально-инерциальной навигации беспилотных летательных аппаратов в условиях недоступности сигналов глобальных навигационных спутниковых систем. Предложен подход к адаптивному формированию ковариации измерений нейросетевого модуля абсолютной визуальной локализации на основе эвиденциального глубокого обучения (Evidential Deep Learning, EDL) с параметризацией нормально-обратным гамма-распределением. Введено понятие степени объяснимой наблюдаемости как контекстно-зависимой количественной характеристики информативности абсолютных измерений. Доказана теорема о вырождении наблюдаемости при стандартной ковариации процесса и предложена её модификация, обеспечивающая ненулевой коэффициент усиления Калмана между обновлениями. Проведено систематическое экспериментальное исследование, включающее четыре формулировки адаптивной ковариации измерений, два режима ковариации процесса и алгоритм автоматического переключения стратегий на двух синтетических и одном полётном наборе данных. Эвиденциальная ковариация превосходит эвристику на всех трёх наборах; величина выигрыша монотонно возрастает с информативностью сцены: снижение средней траекторной ошибки на 7-46% относительно эвристики. На наиболее информативном наборе достигается снижение на 46% и увеличение доли точных кадров с 27% до 66%. Модифицированная ковариация процесса повышает коэффициент усиления Калмана в конфигурациях, где между абсолютными обновлениями накапливается дрейф. Анализ границ применимости показывает, что величина выигрыша определяется корреляцией предсказанной EDL ошибки с фактической (от 0,16 до 0,96 на разных наборах), что задаёт количественный критерий применимости метода.

Ключевые слова: беспилотный летательный аппарат; визуально-инерциальная навигация; фильтр Калмана с ошибкой состояния; абсолютная визуальная локализация; нейросетевой дескриптор; эвиденциальное глубокое обучение; адаптивная ковариация.

Для цитирования: Лазарева П.А., Гурько Н.В. Визуально-инерциальная навигация БПЛА с адаптивной оценкой неопределённости нейросетевой локализации на основе эвиденциального обучения. Труды ИСП РАН, 2026, том 38, вып. 6, часть 1, стр. 259–278. DOI: 10.15514/ISPRAS-2026-38(6)–17.

Благодарности: Работа выполнена при поддержке предоставленного в 2025 году Фондом науки и технологий Республики Татарстан гранта на осуществление фундаментальных и поисковых исследований в научных и образовательных организациях, предприятиях и организациях реального сектора экономики Республики Татарстан.

Visual-inertial navigation of UAVs with adaptive uncertainty estimation of neural network localization based on evidential learning

¹ P.A. Lazareva, ORCID: 0000-0002-1393-3091 <palazareva@kai.ru>
² N.V. Gurko, ORCID: 0009-0001-0217-0471 <nickolay73@hotmail.com>

Kazan National Research Technical University named after A. N. Tupolev – KAI,
10, Karl Marx st., Kazan, 420111, Russia.

Abstract. The problem of unmanned aerial vehicle navigation in GNSS-denied environments is considered. A hybrid system based on an error-state Kalman filter (ESKF) is proposed, integrating inertial measurements, optical-flow velocity estimation, and absolute visual localization against satellite maps using XFeat neural-network descriptors and the LighterGlue matching algorithm. To adaptively form the measurement covariance of the neural localization module, an evidential deep learning (EDL) approach with a Normal-Inverse-Gamma parameterization is proposed, decomposing localization uncertainty into aleatoric and epistemic components in a single forward pass. A notion of explainable observability is introduced as a context-dependent measure of absolute-measurement informativeness, a theorem on observability degeneracy under standard process covariance is proven, and a modification ensuring a non-zero Kalman gain between updates is given. A systematic ablation study over four measurement-covariance formulations, two process-covariance regimes, and an online strategy-switching algorithm was conducted on three flight datasets. The evidential covariance outperforms the heuristic on all three datasets; the gain grows monotonically with scene informativeness, reducing the mean trajectory error by 7–46 % relative to the heuristic. The magnitude of the gain is governed by the correlation between the EDL-predicted error and the actual error (ranging from 0.16 to 0.96 across datasets), which provides a quantitative applicability criterion. The clipped formulation, anchoring the estimate to the heuristic prior, is shown to be a necessary robustness mechanism on domains where the epistemic component of the NIG output is miscalibrated.

Keywords: unmanned aerial vehicle; visual-inertial navigation; error state Kalman filter; absolute visual localization; neural network descriptor; explainable artificial intelligence; deep learning; epistemic uncertainty

For citation: Lazareva P.A., Gurko N.V. Visual-inertial navigation of UAVs with adaptive uncertainty estimation of neural network localization based on evidential learning. Trudy ISP RAN/Proc. ISP RAS, 2026, vol. 38, issue 6, part 1, pp. 259-278 (in Russian). DOI: 10.15514/ISPRAS-2026-38(6)-17.

Acknowledgements. This work was supported by a grant awarded in 2025 by the Science and Technology Fund of the Republic of Tatarstan for the implementation of fundamental and exploratory research in scientific and educational organizations, enterprises, and organizations in the real sector of the economy of the Republic of Tatarstan.

1. Введение

Навигация беспилотных летательных аппаратов (БПЛА) при недоступности глобальных навигационных спутниковых систем (ГНСС) является одной из ключевых задач современной робототехники [1-2]. Основной подход к решению данной проблемы опирается на интеграцию инерциальных измерений с визуальными. В этих условиях система навигации должна обеспечивать три ключевых свойства: глобальную согласованность координат, высокую частоту обновлений и возможность работы на ограниченных бортовых ресурсах. Существующие методы решают эту задачу лишь частично. Визуальная одометрия даёт высокую частоту, но накапливает дрейф. Абсолютная визуальная локализация (АВЛ) даёт глобальные координаты, но работает медленно – порядка одного-пяти герц. Методы одновременной локализации и построения карты ресурсоёмки и формируют только локальные координаты.

Слияние данных от инерциального измерительного блока (ИИБ), оптического потока (ОП) для высокочастотного счисления пути и абсолютной визуальной локализации для коррекции дрейфа в рамках расширенного фильтра Калмана с ошибкой состояния (Error-State Kalman

Filter, ESKF) [3] позволяет объединить преимущества этих подходов и получить стабильные оценки координат и состояния борта с достаточной частотой и абсолютной привязкой.

В последние годы достигнут существенный прогресс в качестве абсолютной локализации за счёт внедрения нейросетевых детекторов и алгоритмов сопоставления признаков, таких как XFeat [4] и LightGlue [5]. Однако переход от классических аналитических измерительных моделей к нейросетевым ставит принципиальную проблему: измерительная функция $h(\mathbf{x}, \mathbf{I}; \boldsymbol{\theta})$, сопоставляющая состояние $\mathbf{x}$ и изображение $\mathbf{I}$ с предсказанным измерением, где $\boldsymbol{\theta}$ – обучаемые веса нейросети, становится обучаемой стохастической, а не детерминированной, и величина ковариации шума измерений $\mathbf{R}_k$ зависит от контекста визуальной сцены, качества сопоставления и степени доменного разрыва между текущим изображением и обучающей выборкой. Использование фиксированной или эвристической $\mathbf{R}_k$ – стандартный подход в существующих работах [6] – приводит к двум разнонаправленным проблемам: на простых, информативных сценах фильтр недостаточно доверяет точному измерению (заниженный коэффициент усиления Калмана), тогда как на сложных, текстурно-бедных сценах фильтр переоценивает его точность, что приводит к расхождению. Адаптивная ковариация измерений требуется для разрешения этого противоречия.

Существующие подходы к адаптации $\mathbf{R}_k$ в нейросетевой навигации можно разделить на три группы. Первая – эвристическая адаптация на основе числа согласованных точек $\mathbf{R}_k = \sigma^2 \text{heur}(n_{\text{inl}}) \mathbf{I}$ [7] – проста и работоспособна, но не различает геометрически и семантически различные ситуации с одинаковым n_{inl} . Вторая – байесовские нейронные сети и прореживание Монте-Карло [8-9] – обеспечивает теоретически обоснованные оценки неопределённости, но требует T -кратного увеличения вычислений и не пригодна для бортовых систем реального времени. Третья – методы, обучающие отдельную нейронную сеть для предсказания ковариации шума процесса [10-11] – повышает робастность, но не решает задачу адаптации измерительной ковариации в зависимости от семантики сцены.

Эвиденциальное глубокое обучение (evidential deep learning, EDL) [12] предлагает альтернативу, объединяющую преимущества байесовских методов с однопроходной вычислительной эффективностью. Параметризуя выход нейросети нормально-обратным гамма-распределением (Normal-Inverse-Gamma, NIG), EDL разделяет алеаторную и эпистемическую компоненты неопределённости за один прямой проход через сеть: алеаторная неопределённость отражает неустранимую неопределённость, обусловленную шумом и неоднозначностью входных данных, тогда как эпистемическая связана с недостаточностью знаний модели и ограниченностью обучающей выборки. В задачах локализации по сценам этот подход применялся для построения карт информативности и планирования траектории [13], но его интеграция в реальный конвейер ESKF + оптический поток + абсолютная локализация и теоретический анализ влияния на наблюдаемость до настоящего времени не рассматривались.

Настоящая работа предлагает следующие основные результаты:

1. Формализация объяснимой наблюдаемости в визуально-инерциальной навигации. Введена скалярная контекстно-зависимая степень XO_k , обобщающая бинарный ранговый критерий классической теории наблюдаемости и непосредственно связанная с коэффициентом усиления Калмана.
2. Теорема о вырождении наблюдаемости при стандартной ковариации процесса (Теорема и её устранение через модифицированную ковариацию $\mathbf{Q}_p^{\text{mod}}$, обеспечивающую линейный рост позиционной ковариации между абсолютными обновлениями.

3. Систематическое экспериментальное исследование четырёх формулировок адаптивной ковариации $\mathbf{R}_k$ в сочетании с двумя режимами $\mathbf{Q}_k$ на двух синтетических и одном полётном наборах данных. Проведено 24 воспроизводимых прогона с фиксированным детерминированием; обеспечена бит-точная воспроизводимость результатов, то есть полное совпадение выходных данных при повторных запусках вплоть до отдельных битов их цифрового представления.
4. Экспериментальное подтверждение Теоремы о вырождении: модифицированная ковариация процесса повышает коэффициент усиления Калмана на 22 – 26% в конфигурациях, где между абсолютными обновлениями накапливается дрейф.

Анализ границ применимости. На наборе с информативной структурой сцены сочетание EDL-ковариации с модифицированной $\mathbf{Q}_p^{\text{mod}}$ обеспечивает снижение средней траекторной ошибки на 44,3% относительно эвристики при увеличении доли точек с малой ошибкой (< 5 м) в 2,9 раза. Эвиденциальная ковариация превосходит эвристику на всех трёх наборах; величина выигрыша монотонно возрастает с информативностью сцены (от 7 % на наборе с ровным полётом без манёвров до 46 % на наиболее информативном наборе) и количественно определяется корреляцией предсказанной EDL ошибки с фактической (от 0,16 до 0,96, соответственно) – это задаёт количественный критерий применимости метода.

Статья организована следующим образом. В разделе 2 кратко рассматривается состояние области. В разделе 3 излагается теоретическая основа: динамическая модель, информационная характеристика наблюдаемости, NIG-параметризация неопределённости, теорема о вырождении и её устранение. В разделе 4 описана архитектура системы. В разделе 5 представлены экспериментальные результаты, включая верификацию теоремы и анализ границ применимости. Раздел 6 содержит обсуждение и направления дальнейшей работы.

2. Связанные работы

Современные подходы к визуально-инерциальной навигации (ВИН) базируются либо на оптимизации [14-15] (VINS-Mono, OKVIS), либо на фильтрационном подходе [3, 16] (MSCKF, ROVIO). Фильтрационные методы предпочтительнее для бортовых платформ из-за низких вычислительных требований; в данной работе используется ESKF-формулировка с разделением номинального состояния и оценок ошибок [3]. Все указанные методы предполагают известную ковариацию измерений; адаптация $\mathbf{R}_k$ к контексту сцены остаётся открытой проблемой.

Нейросетевая абсолютная локализация по карте основана на сопоставлении бортового изображения с предварительно загруженной геопривязанной картой. В последние годы доминирующим стал подход на основе нейросетевых детекторов разреженных признаков и обучаемых алгоритмов сопоставления: SuperPoint+SuperGlue [17], DISK+LightGlue [5], XFeat+LightGlue [4-5]. В работе [18] продемонстрировано применение XFeat+LightGlue с семантическими дескрипторами для навигации БПЛА в условиях отсутствия сигналов ГНСС. В [19] предложен метод сопоставления изображений с разрешением проблем масштаба и перспективы. Однако во всех указанных работах ковариация выходных позиционных измерений либо считается фиксированной, либо моделируется простой эвристикой от числа согласованных точек.

Классические подходы к адаптивной ковариации в фильтрах Калмана основаны на статистике невязок [20]. Современные нейросетевые расширения этого подхода включают архитектуру ProcessNet в связке с адаптивным фильтром Калмана (ANUKF) [10] для подводной навигации, гибридный нейросетевой сигма-точечный фильтр Калмана для наземных платформ [11] и гибридную архитектуру Bayesian KalmanNet [21]. В большинстве

этих работ адаптируется ковариация процесса; адаптация измерительной ковариации в задаче нейросетевой визуальной локализации систематически не рассматривалась.

Подход эвиденциального обучения предложен в [12] на основе NIG-параметризации, обеспечивающей разделение алеаторной и эпистемической неопределённости за один прямой проход. В [13] EDL применено в задаче регрессии координат сцены для целей перцептивно-ориентированного планирования траектории БПЛА. В [22] показано использование EDL для регрессии эмоций; в [23] – для прогнозирования временных рядов. Применение EDL для адаптации измерительной ковариации в реальном вычислительном конвейере ESKF + ОП + АВЛ в задаче ВИН до настоящего времени, насколько нам известно, не предлагалось.

В работе [24] продемонстрировано, что включение схемы адаптивного масштабирования измерительной ковариации на основе теплокарт нейросети существенно повышает робастность фильтра в задаче определения положения космического аппарата. Этот результат является методологическим прецедентом для предлагаемого подхода в задаче визуально-инерциальной навигации БПЛА.

В отличие от перечисленных работ, в настоящей статье впервые предложена интеграция эвиденциальной оценки неопределённости в полный конвейер ВИН с теоретическим обоснованием через расширение классической теории наблюдаемости и экспериментальной верификацией ключевых теоретических положений на воспроизводимых прогонах.

3. Теоретическая основа

3.1 Динамическая модель и каналы измерений

Рассмотрим нелинейную дискретную динамическую систему, описывающую движение БПЛА:

$$\mathbf{x}_{k+1} = f(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{G}_k \mathbf{w}_k, \quad \mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k), \quad (1)$$

где $\mathbf{x}_k \in \mathbb{R}^{16}$ – вектор состояния, включающий положение $\mathbf{p} \in \mathbb{R}^3$, скорость $\mathbf{v} \in \mathbb{R}^3$, кватернион ориентации $\mathbf{q} \in S^3$, смещения акселерометра $\mathbf{b}_a \in \mathbb{R}^3$ и гироскопа $\mathbf{b}_g \in \mathbb{R}^3$; $\mathbf{u}_k$ – измерения инерциального измерительного блока; $\mathbf{w}_k$ – белый гауссов шум процесса с ковариацией $\mathbf{Q}_k$; $\mathbf{G}_k$ – матрица влияния шума.

Система располагает двумя каналами измерений различной природы.

Канал 1 (высокочастотный) передает скоростные измерения по оптическому потоку:

$$\mathbf{z}_k^{(v)} = \mathbf{H}_v \mathbf{x}_k + \mathbf{n}_k^{(v)}, \quad \mathbf{n}_k^{(v)} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_v), \quad (2)$$

где $\mathbf{H}_v = [\mathbf{0}_{2 \times 3} \quad \mathbf{I}_2 \quad \mathbf{0}_{2 \times 11}]$ – линейная матрица наблюдения, извлекающая горизонтальные компоненты скорости, $\mathbf{n}_k^{(v)}$ – белый гауссов шум оптического потока с ковариацией $\mathbf{R}_v$.

Канал 2 (низкочастотный) передает абсолютные позиционные измерения по визуальной локализации:

$$\mathbf{z}_k^{(p)} = h(\mathbf{x}_k, \mathbf{I}_k; \mathbf{\theta}) + \mathbf{n}_k^{(p)}. \quad (3)$$

где $\mathbf{I}_k$ – изображение, полученное камерой; $\mathbf{\theta}$ – параметры нейросетевой модели сопоставления; $h: \mathbb{R}^n \times \mathcal{I} \rightarrow \mathbb{R}^3$ – нелинейное отображение, определяющее положение БПЛА по результатам сопоставления изображения с геопривязанной картой.

Ключевое отличие канала 2 от канала 1 состоит в том, что измерительное отображение h является обучаемым стохастическим, а не аналитическим детерминированным. Это

порождает неопределённость, не сводимую к аддитивному гауссову шуму $\mathbf{n}_k^{(p)}$ с фиксированной ковариацией.

3.2 Информационная характеристика наблюдаемости

Классическое определение локальной наблюдаемости через ранг матрицы наблюдаемости [25] неадекватно в случае нейросетевых измерительных модулей по трём причинам:

- 1) ранг может формально сохраняться, но фактическая информативность измерений варьируется на порядки в зависимости от визуальной сцены;
- 2) ковариация $\mathbf{R}_k^{(p)}$ является контекстно-зависимой переменной, а не константой;
- 3) неопределённость модели h содержит как алеаторную, так и эпистемическую компоненты, по-разному влияющие на наблюдаемость.

Для перехода от бинарной к количественной характеристике введём информационный формализм.

Определение 3.1 (Кумулятивная информационная матрица). Для системы (1)–(3) с двумя каналами измерений кумулятивная информация Фишера за один шаг фильтрации:

$$\mathcal{I}_k^{\text{cum}} = \mathbf{H}_v^\top \mathbf{R}_v^{-1} \mathbf{H}_v + \sum_{j \in \mathcal{T}_p} j \in \mathcal{T}_p \mathbf{H}_p^\top \left(\mathbf{R}_j^{(p)} \right)^{-1} \mathbf{H}_p, \quad (4)$$

где $\mathcal{T}_p$ – множество моментов поступления абсолютных измерений; $\mathbf{H}_p = [\mathbf{I}_3 \quad \mathbf{0}_{3 \times 13}]$ – матрица наблюдения положения.

Из границы Крамера–Рао для апостериорной оценки $\hat{\mathbf{x}}_k$ следует, что нижняя граница среднеквадратичной ошибки положения определяется обратной ковариацией абсолютных измерений $\left(\mathbf{R}_k^{(p)} \right)^{-1}$. При фиксированной $\mathbf{R}_0$ эта граница не зависит от визуальной сцены, что является принципиальным ограничением: одинаковое число согласованных точек на текстурированной застройке и на однородном поле даёт одинаковую $\mathbf{R}_0$, хотя фактическая информативность принципиально различна.

Определение 3.2 (Степень наблюдаемости по Фишеру). Скалярная степень наблюдаемости состояния $\mathbf{x}_k$ по каналу абсолютной локализации:

$$\alpha_k^{(p)} = \frac{\text{tr}(\mathbf{H}_p^\top (\mathbf{R}_k^{(p)})^{-1} \mathbf{H}_p)}{\text{tr}((\mathbf{P}_k^-)^{-1})} = \frac{\text{tr}((\mathbf{R}_k^{(p)})^{-1})}{\text{tr}((\mathbf{P}_k^-)^{-1})} .. \quad (5)$$

Величина $\alpha_k^{(p)} \in [0, \infty)$ характеризует относительный вклад абсолютного измерения в информацию о состоянии: при $\alpha_k^{(p)} \gg 1$ измерение доминирует над априорной информацией, при $\alpha_k^{(p)} \ll 1$ оно несущественно, при $\alpha_k^{(p)} \rightarrow 0$ происходит вырождение наблюдаемости.

Утверждение 3.1 (Связь степени наблюдаемости с коэффициентом усиления). В скалярном приближении для одной компоненты положения справедливо:

$$K_k = \frac{\alpha_k^{(p)}}{\alpha_k^{(p)} + 1}. \quad (6)$$

Доказательство. Из формулы Калмана $K = P / (P + R) = (1 / R) / (1 / P + 1 / R) = \alpha / (1 + \alpha)$, где $\alpha = (1 / R) / (1 / P)$.

Из (6) следует, что при $\alpha_k^{(p)} \approx 0,01$ коэффициент усиления $K_k \approx 0,01$, и абсолютные поправки фактически игнорируются. Этот эффект составляет проблему вырождения коэффициента усиления Калмана, формализуемую ниже.

3.3 NIG-параметризация неопределённости

Для аппроксимации неизвестной функции $\mathbf{R}_k(\xi)$, где ξ_k – контекстный вектор измерения (число согласованных точек, остатки, характеристики текстуры и т. д.), используем подход эвиденциального глубокого обучения [12], основанный на нормально-обратном гамма распределении.

Определение 3.3 (NIG-параметризация неопределённости). Обучаемая модель $\theta: \mathbb{R}^d \rightarrow \mathbb{R}^4$ отображает контекстный вектор ξ_k в параметры NIG-распределения:

$$\theta(\xi_k) = (\gamma_k, \nu_k, \alpha_k, \beta_k), \quad \nu_k > 0, \alpha_k > 1, \beta_k > 0, \quad (7)$$

где γ_k – предсказание ожидаемой ошибки измерения; ν_k – число виртуальных наблюдений (уверенность в среднем); α_k, β_k – параметры формы и масштаба обратного гамма-распределения для дисперсии. NIG-распределение является сопряжённым априорным для гауссовой модели с неизвестным средним и дисперсией.

Определение 3.4 (Алеаторная и эпистемическая неопределённости). Для NIG-параметризации (7):

$$\sigma_{\text{alea}}^2(\xi_k) = \frac{\beta_k}{\alpha_k - 1}, \quad \sigma_{\text{epist}}^2(\xi_k) = \frac{\beta_k}{\nu_k(\alpha_k - 1)}. \quad (8)$$

Алеаторная компонента σ_{alea}^2 отражает неустранимую неопределённость, обусловленную разрешением карты, шумом камеры и геометрией сцены, и не снижается при увеличении объёма обучающих данных. Эпистемическая компонента σ_{epist}^2 отражает неопределённость модели, обусловленную ограниченностью обучающей выборки, и снижается при $\nu_k \rightarrow \infty$.

3.4 Объяснимая наблюдаемость и адаптивная ковариация

Определение 3.5 (Адаптивная ковариация измерений). Ковариация абсолютного позиционного измерения, формируемая на основе эвиденциальной оценки:

$$\mathbf{R}_k^{(p)} = \left(\sigma_{\text{alea}}^2(\xi_k) + \sigma_{\text{epist}}^2(\xi_k) \right) \cdot \mathbf{I}_3 = \sigma_{\text{total}}^2(\xi_k) \cdot \mathbf{I}_3. \quad (9)$$

Замечание. Целевая переменная $y = \|e\|_2$ в (7) распределена не нормально, а по Рэлею при допущении $e \sim \mathcal{N}(\mathbf{0}, \sigma_{\text{pos}}^2 \mathbf{I})$. В этом случае предсказанная величина γ интерпретируется как $E[\|e\|]$, что связано с σ_{pos} соотношением $\gamma = \sigma_{\text{pos}} \cdot \sqrt{(\pi/2)}$. Формулировка *rayleigh* (см. раздел 5.3) реализует это преобразование явно, тогда как (9) использует σ_{alea} как непосредственную оценку σ_{pos} , что соответствует ослабленному допущению о гауссовости остатков и обеспечивает консервативную (заниженную) оценку. Сравнение этих формулировок в разделе 5 эмпирически выявляет, какое из приближений лучше согласовано с физикой шума локализации в данном домене.

Определение 3.6 (Степень объяснимой наблюдаемости). Скалярная степень объяснимой наблюдаемости по каналу абсолютной локализации:

$$XO_k = \frac{3 / \sigma_{\text{total}}^2(\xi_k)}{\text{tr}(\mathbf{P}_k^-)^{-1} [1:3, 1:3]}. \quad (10)$$

В отличие от классической степени наблюдаемости $\alpha_k^{(p)}$ из (5), XO_k является контекстно-зависимой: она варьируется от кадра к кадру в зависимости от визуальной сцены, качества сопоставления и уверенности модели.

Условие калибровки. Модель θ называется калиброванной на уровне $1 - \delta$, если для предсказанных доверительных интервалов выполняется:

$$\Pr(\|z_k^{(p)} - \mathbf{p}_k\| \leq c_\delta \cdot \sigma_{\text{total}}(\xi_k)) \geq 1 - \delta, \quad \forall k. \quad (11)$$

3.5 Теорема об информационном выигрыше

Теорема 3.1 (Информационный выигрыш адаптивной ковариации). Пусть $\mathbf{R}_0 = \sigma_0^2 \mathbf{I}_3$ – фиксированная ковариация (эвристика), а $\mathbf{R}_k^{\text{adapt}} = \sigma_{\text{total}}^2(\xi_k) \mathbf{I}_3$ – адаптивная ковариация (9). Пусть модель θ удовлетворяет условию калибровки (11) на уровне $1 - \delta$ при $\sigma_0 = \max_k \sigma_{\text{total}}(\xi_k)$. Тогда кумулятивная информация Фишера за K шагов при адаптивной ковариации не меньше, чем при фиксированной:

$$\sum_{k=1}^K \frac{1}{\sigma_{\text{total}}^2(\xi_k)} \geq \sum_{k=1}^K k = 1^K \frac{1}{\sigma_0^2}. \quad (12)$$

Доказательство. Для обеспечения покрытия (11) при фиксированной ковариации необходимо $\sigma_0 \geq \max_k \sigma_{\text{total}}(\xi_k)$. Тогда для всех k выполняется $\sigma_{\text{total}}(\xi_k) \leq \sigma_0$, и строго меньше для информативных измерений. Следовательно, $1 / \sigma_{\text{total}}^2(\xi_k) \geq 1 / \sigma_0^2$ для всех k , откуда (12).

Содержательный смысл: адаптивная ковариация извлекает больше информации из «хороших» кадров (текстурированная застройка, яркое освещение), не жертвуя робастностью на «плохих» (однородное поле, размытие).

3.6 Теорема о вырождении наблюдаемости

В реальной системе с двумя каналами измерений возникает следующая проблема: высокочастотные скоростные обновления оптического потока сжимают позиционную ковариацию $\mathbf{P}_{pp}$ до значений, при которых низкочастотные абсолютные обновления получают практически нулевой коэффициент усиления.

Теорема 3.2 (Условие вырождения и его устранение). Рассмотрим ESKF с шумом процесса $\mathbf{Q}_k$ вида (1) и двумя каналами измерений (2)–(3). Пусть скоростные обновления поступают с частотой f_v , а позиционные – с частотой $f_p \ll f_v$. Тогда:

- (а) Вырождение. При стандартной ковариации процесса $\mathbf{Q}_p^{\text{std}} = \frac{1}{4} \sigma_a^2 t^4 \mathbf{I}_3$ после $N = f_v / f_p$ скоростных обновлений без позиционных:

$$\text{tr}(\mathbf{P}_{pp}^-) \rightarrow \text{tr}(\mathbf{P}_{pp}, 0) + \frac{1}{4} \sigma_a^2 t^4 N \approx \text{tr}(\mathbf{P}_{pp}, 0), \quad (13)$$

поскольку $\sigma_a^2 t^4 \ll \text{tr}(\mathbf{P}_{pp}, 0)$ для типичных параметров. Скоростные обновления сжимают $\mathbf{P}_{pp}$ до значений порядка $R_v / (\mathbf{H}_v \mathbf{H}_v^\top)$, и $\alpha_k^{(p)} \rightarrow 0$.

(б) Устранение. При модифицированной ковариации:

$$\mathbf{Q}_p^{\text{mod}} = \max\left(\frac{1}{4}\sigma_a^2 t^4, 2\sigma_p^2 \tau_p\right) \mathbf{I}_3 \quad (14)$$

после N шагов без позиционного обновления:

$$\text{tr}(\mathbf{P}_{pp}^-) \rightarrow 2\sigma_p^2 \tau_p \cdot \frac{N}{f_p} = \frac{2\sigma_p^2 \tau_p \cdot 3}{f_p}, \quad (15)$$

что обеспечивает $\alpha_k^{(p)} \geq 3 / (f_p \mathbf{R}_k^{(p)})$, то есть ненулевую степень наблюдаемости при любой $\mathbf{R}_k^{(p)}$.

Доказательство. Часть (а) следует из рекуррентного применения формулы обновления ковариации (форма Джозефа) с $\mathbf{R}_v$ порядка единиц м/с² и $\mathbf{Q}_p^{\text{std}}$ порядка 10^{-8} м². После одного скоростного обновления $\mathbf{P}_{pp}^+ = (\mathbf{I} - \mathbf{K}_v \mathbf{H}_v) \mathbf{P}_{pp}^-$, и поскольку $\mathbf{H}_v$ затрагивает только скоростные компоненты, а $\mathbf{P}_{pp}$ связана со скоростью через $\mathbf{P}_{pv}$, эффективное сжатие $\mathbf{P}_{pp}$ при последовательных скоростных обновлениях определяется недиагональными элементами и стремится к асимптотическому значению порядка $\mathbf{R}_v / \mathbf{H}_v \mathbf{H}_v^\top$. Слагаемое $\frac{1}{4}\sigma_a^2 t^4$ за N шагов накапливает прирост порядка $\sigma_a^2 t^4 N$, что для типичных значений $\sigma_a \sim 10^{-3}$ м/с² и $t \sim 0,1$ с пренебрежимо мало по сравнению с инициализационным $\text{tr}(\mathbf{P}_{pp,0})$ и асимптотическим уровнем после скоростных обновлений.

Часть (б) следует из того, что слагаемое $2\sigma_p^2 \tau_p \mathbf{I}_3$ обеспечивает линейный рост $\mathbf{P}_{pp}$ между позиционными обновлениями со скоростью $2\sigma_p^2 \tau_p$ за единицу времени. За интервал $T = 1 / f_p$ между АВЛ-обновлениями накапливается прирост $2\sigma_p^2 \tau_p / f_p$ на каждую из трёх диагональных компонент, откуда (15). Подставляя в (5), получаем нижнюю оценку для $\alpha_k^{(p)}$.

Параметры σ_p и τ_p выбираются исходя из физической природы шума процесса в позиционном канале и связаны с типичной частотой АВЛ-обновлений: $\sigma_p \approx \sigma_{\text{AVL}}$, $\tau_p \approx 1 / f_p$.

3.7 Согласованность фильтра

Определение 3.7 ($\mathcal{E}$ -согласованность фильтра). Фильтр называется $\mathcal{E}$ -согласованным, если для нормализованной невязки $\mathbf{y}_k = \mathbf{z}_k - \mathbf{H} \hat{\mathbf{x}}_k^-$ выполняется:

$$\mathbb{E}[\mathbf{y}_k^\top \mathbf{S}_k^{-1} \mathbf{y}_k] = \text{dim}(\mathbf{y}) \pm O(\varepsilon), \quad (16)$$

где $\mathbf{S}_k = \mathbf{H} \mathbf{P}_k^- \mathbf{H}^\top + \mathbf{R}_k$ — ковариация невязки.

Теорема 3.3 (Согласованность при калиброванной ковариации). Если модель $\mathbf{0}$ удовлетворяет условию калибровки (11) с уровнем $1 - \delta$, и шум процесса $\mathbf{w}_k$ гауссов с известной ковариацией $\mathbf{Q}_k$, то ESKF с адаптивной ковариацией $\mathbf{R}_k^{(p)} = \sigma_{\text{total}}^2(\xi_k) \mathbf{I}_3$ является $\mathcal{E}$ -согласованным с $\varepsilon = O(\delta)$.

Доказательство. Согласованность фильтра Калмана эквивалентна тому, что нормализованная невязка $\mathbf{S}_k^{-1/2} \mathbf{y}_k$ распределена как стандартная нормальная. Если $\mathbf{R}_k^{(p)}$ калибрована в смысле

(11), то $\mathbf{R}_k^{(p)} \approx \text{cov}(\mathbf{n}_k^{(p)})$ с вероятностью $1 - \delta$, что гарантирует $\text{cov}(\mathbf{y}_k) \approx \mathbf{S}_k$. Формальный вывод следует из обобщения [26] на случай адаптивных ковариаций.

4. Архитектура системы

Предлагаемая система визуально-инерциальной навигации (рис. 1) объединяет три источника информации: инерциальные измерения от ИИБ, оптический поток между последовательными кадрами бортовой камеры, направленной в надир, и абсолютную визуальную локализацию через сопоставление текущего изображения с предварительно построенной картой признаков. Все три потока интегрируются в едином ESKF, ковариация измерений канала абсолютной локализации формируется адаптивно по выходам нейросетевого модуля EDL.

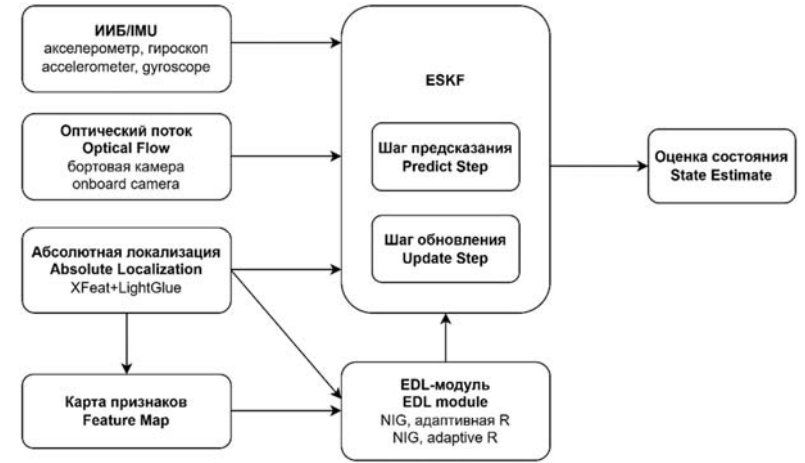


Рис. 1. Архитектура системы визуально-инерциальной навигации с EDL-модулем.
Fig. 1. Architecture of the visual-inertial navigation system with the EDL module.

4.1 ESKF с двумя каналами измерений

Состояние фильтра включает положение, скорость, кватернион ориентации и смещения акселерометра и гироскопа, $\text{dim}(\mathbf{x}) = 16$. Шаг предсказания выполняется на каждом отсчёте измерений ИИБ (~ 100 Гц) по нелинейной модели механизации, ковариация распространяется через линеаризованный якобиан $\mathbf{F}_k$:

$$\hat{\mathbf{x}}_{k+1} = f(\hat{\mathbf{x}}_k^+, \mathbf{u}_k), \quad \mathbf{P}_{k+1}^- = \mathbf{F}_k \mathbf{P}_k^+ \mathbf{F}_k^\top + \mathbf{Q}_k. \quad (17)$$

Шаги обновления для каналов скорости и положения выполняются в стандартной форме Джозефа [27], при необходимости предварительно проходят проверку допустимости по расстоянию Махаланобиса. Полные формулы механизации и обновлений ESKF приводятся в [28]; ниже описаны только специфические для предлагаемого подхода компоненты.

Модифицированная ковариация процесса. В соответствии с Теоремой 3.2(b), к стандартной диагональной ковариации позиционных компонент добавляется слагаемое, обеспечивающее линейный рост $\mathbf{P}_{pp}$ между АВЛ-обновлениями:

$$\mathbf{Q}_p^{\text{mod}} = \mathbf{Q}_p^{\text{std}} + 2\sigma_p^2 \tau_p \cdot \Delta t \cdot \mathbf{I}_3, \quad (18)$$

где Δt – интервал между шагами предсказания. В экспериментах принято $\sigma_p = 5$ м, $\tau_p = 1$ с. Эти значения выбраны исходя из типичной шкалы шума АВЛ (~3-5 м) и характерного интервала между АВЛ-обновлениями (~0,2 с).

4.2 Оптический поток

Оценка скорости через оптический поток выполняется на разреженном наборе угловых точек, детектируемых по критерию Ши–Томэзи с равномерным распределением по сетке 8×6 . Сопоставление между парами кадров производится методом Лукаса–Канаде с обратной проверкой, что отбраковывает ненадёжные точки. Геометрическая модель – частичное аффинное преобразование, оценивается методом RANSAC. Скорость восстанавливается из смещения центра изображения с компенсацией изменения углов крена и тангажа:

$$\Delta x_{att} = f_x \cdot (\tan \phi_k - \tan \phi_{k-1}), \quad \Delta y_{att} = f_y \cdot (\tan \theta_k - \tan \theta_{k-1}), \quad (19)$$

где ϕ, θ – углы крена и тангажа, f_x, f_y – фокусные расстояния. Скоростные компоненты в навигационной системе координат получаются после трёхосного поворота с учётом высоты и углов отношения. Ковариация $\mathbf{R}_v$ оценивается из медианного отклонения остатков аффинной модели.

4.3 Абсолютная визуальная локализация

Абсолютные позиционные измерения формируются сопоставлением текущего бортового изображения с предварительно построенной картой признаков геопривязанного спутникового снимка. Используются нейросетевые архитектуры XFeat [4] для детектирования и описания локальных признаков и LightGlue [5] для их сопоставления, соответственно. На этапе предполётной подготовки осуществляется пространственная декомпозиция карты на фрагменты, для каждого фрагмента извлекаются признаки с числом ключевых точек до 8192. На этапе локализации:

1. из бортового изображения извлекается до 1024 ключевых точек;
2. из карты выбирается фрагмент размером 1500×1500 пикселей вокруг предсказанной позиции;
3. сопоставление производится через LightGlue с порогом 0,45;
4. гомография оценивается через RANSAC с порогом репроекции 8 пикселей;
5. положение в системе координат карты восстанавливается из преобразования центра бортового изображения.

Результирующее позиционное измерение $\mathbf{z}_k^{(p)} = (E_k, N_k, U_k)$ передаётся в ESKF с задержкой, соответствующей времени обработки изображения; в формуле обновления используется экстраполяция по скорости на момент применения коррекции.

4.4 EDL-модуль и контекстный вектор

В реализованной системе EDL-модель θ выполнена в виде компактного многослойного перцептрона (около 1500 параметров), принимающего на вход контекстный вектор $\xi_k \in \mathbb{R}^{10}$:

$$\xi_k = (n_{inl}, n_{raw}, r_{inl}, \sin \psi, \cos \psi, \dot{\psi}, v, \phi, h, s_{render})^\top. \quad (20)$$

где n_{inl} – число согласованных точек RANSAC; n_{raw} – число первичных соответствий; $r_{inl} = n_{inl} / n_{raw}$ – доля согласованных точек; ψ – угол курса (использующий синус-

косинусную циркулярную кодировку); $\dot{\psi}$ – скорость изменения курса; v – скорость полёта; ϕ – крен; h – высота; s_{render} – масштабирующий коэффициент рендеринга.

Выход модели – параметры NIG-распределения $(\gamma_k, \nu_k, \alpha_k, \beta_k)$ с функцией активации softmax для обеспечения положительности: $\nu = \text{softplus}(\nu')$, $\alpha = \text{softplus}(\alpha') + 1$, $\beta = \text{softplus}(\beta')$. Обучение производится минимизацией отрицательного логарифма правдоподобия Стюдента (выводимое из NIG) с регуляризацией по эвиденции, как в [12]. В качестве целевой переменной используется фактическая ошибка локализации в метрах относительно эталонной позиции, известной из ГНСС-разметки тренировочного набора. Сформированная адаптивная ковариация измерений (9) подставляется в стандартную формулу обновления ESKF на каждом успешном АВЛ-измерении.

4.5 Программная реализация

Для обеспечения воспроизводимости результатов реализация выполнена в детерминированном однопоточном режиме: вместо асинхронной обработки через два параллельных потока (быстрый ИИБ/ОП, медленный АВЛ) применён синхронный последовательный конвейер с фиксированной задержкой применения АВЛ-измерений (3 шага предсказания, имитирующих реальную задержку обработки изображения порядка 0,3 с). Для обеспечения воспроизводимости начальные состояния генераторов псевдослучайных чисел библиотек NumPy, PyTorch и OpenCV фиксировались одним и тем же заданным значением (seed). Это обеспечивает бит-точную воспроизводимость траекторных ошибок при повторных прогонах одной и той же конфигурации, что является принципиальным требованием для систематического абляционного исследования.

5. Экспериментальное исследование

5.1 Наборы данных

Для систематической оценки предложенного подхода использованы два синтетических и один полётный наборы данных, существенно различающиеся по информативности визуальной сцены. Использованные участки спутниковых карт приведены на рис. 2.



Рис. 2. Участки спутниковых карт, использованные в наборах данных D1-D3.
Fig. 2. Satellite map tiles used in datasets D1-D3.

Синтетический набор данных D1 – полёт на высоте 900 м длительностью 276 с, 2680 кадров. Карта района – спутниковый снимок размером 10980×10980 пикселей, разрешение карты ≈ 10 м/пкс. Сцена характеризуется высокой текстурной информативностью: яркие зелёные сельскохозяйственные поля с чёткими границами, контуры рек, разнообразные ландшафтные ориентиры.

Набор полётных данных D2 – полёт на высоте 900 м длительностью 290 с, 2897 кадров. Карта района сопоставима по размеру с D1. Сцена характеризуется пониженной

информативностью: преобладание серых и тёмно-зелёных тонов, регулярные структуры (тепличные хозяйства, прямолинейные участки полей), повторяющиеся текстуры, повышающие вероятность неоднозначного или ошибочного сопоставления визуальных признаков. Эталонная траектория восстановлена по данным ГНСС.

Синтетический набор данных D3 представляет низковысотный полёт (275–300 м) с разрешением около 0,6 м/пиксель, выраженными манёврами и повышенным шумом инерциальных измерений. Сцена характеризуется средней информативностью с богатой текстурной структурой.

Различие в информативности сцен между тремя наборами данных служит экспериментальным базисом для определения границ применимости предложенного метода.

5.2 Обучение EDL-модели

Обучающие данные для EDL формируются прогоном алгоритма абсолютной локализации на полётах из соответствующего набора данных с известной эталонной траекторией; для каждого успешного матча в обучающую выборку записывается контекстный вектор ξ_k (20) и фактическая ошибка локализации в метрах. Размещение области интереса для сопоставления выполняется по эталонной позиции, что обеспечивает чистоту обучения; в режиме рабочего вычислительного конвейера область интереса размещается по предсказанной ESKF позиции, что вносит небольшой доменный сдвиг (см. обсуждение в разделе 6).

Архитектура модели – трёхслойный перцептрон с примерно 1500 параметрами и NIG-выходом. Обучение производится с функцией потерь, состоящей из отрицательного логарифма Стьюдент-правдоподобия и эвиденциальной регуляризации [12]. Для каждого набора данных обучается отдельная модель, соответствующая домену.

Характеристики обученных моделей приведены в табл. 1.

Табл. 1. Характеристики EDL-моделей на трёх наборах данных.

Table 1. Characteristics of EDL models on three datasets.

Метрика	D1	D2	D3
$\text{corr}(\gamma, \text{err})$	0,96	0,16	0,56
$\text{corr}(\sigma_{\text{alea}}, \text{err})$	−0,25	0,25	0,45
Калибровка $1\sigma_{\text{total}}$	91,5 %	99,4 %	76 %
Калибровка 99,9 %	100 %	91,8%	89 %
$\bar{\sigma}_{\text{alea}}$ (м)	0,40	4,29	0,46
$\bar{\sigma}_{\text{epist}}$ (м)	6,77	8,83	0,43
$\bar{\gamma}$ (м)	1,60	2,78	0,69

Различие в значениях $\text{corr}(\gamma, \text{err})$ служит ключевым индикатором того, насколько информативен EDL-сигнал в данном домене.

5.3 Конфигурации R_k и Q_k

Для систематического сравнения исследованы четыре формулировки адаптивной ковариации измерений и два режима ковариации процесса.

Формулировки $R_k^{(p)}$:

- эвристическая (heuristic): $\sigma_k = \max(3, 50 / \sqrt{n_{\text{ini}}})$. Базовый подход без использования EDL;
- рэлеевская (rayleigh): $\sigma_k = \max(3, \gamma_k / \sqrt{\pi/2})$. Используется только γ , преобразование через распределение Рэлея учитывает двумерность остаточного вектора (E, N) ;
- теоретическая формула (theoretical): $\sigma_k^2 = \sigma_{\text{alea}}^2 + \sigma_{\text{epist}}^2 = \sigma_{\text{total}}^2$. Прямая реализация Определения 3.5;
- теоретическая с зажимом (theoretical_clipped): $\sigma_k^2 = \sigma_{\text{alea}}^2 \cdot \text{clip}(\sigma_{\text{alea}} / \sigma_{\text{heur}}, 0, 7, 1, 3)$. Инженерная регуляризация против выбросов σ_{epist} .

Режимы Q_k :

- stdQ: стандартная диагональная ковариация процесса;
- modQ: модифицированная ковариация (18) с дополнительным позиционным слагаемым.

Итого $4 \times 2 = 8$ конфигураций на каждом наборе данных, 24 прогона суммарно по трём наборам.

5.4 Метрики и протокол прогонов

Основная метрика – абсолютная траекторная ошибка (Absolute Trajectory Error, ATE), определённая как евклидово расстояние между оценкой положения ESKF и эталонной траекторией на каждом кадре. Отчитываются: среднее (Mean ATE), медиана (Median ATE), доля кадров с ошибкой не более 5 м (< 5 м, %) и доля кадров с ошибкой не более 10 м (< 10 м, %).

Дополнительные диагностические метрики:

- эффективный коэффициент усиления Калмана K_{axis} – оценивается по падению $\text{tr}(\mathbf{P}_{pp})$ при АВЛ-обновлениях: $K \approx -\Delta \text{tr}(\mathbf{P}_{pp}) / \text{tr}(\mathbf{P}_{pp})^-$;
- калибровка адаптивной ковариации – доля кадров, в которых $|\text{err}| \leq c\sigma_{\text{used}}$, для $c \in 1, 2, 3$;
- корреляция предсказанной ошибки γ с фактической.

Все прогоны выполнены в детерминированном однопоточном режиме с фиксированным начальным значением, равным 42, что обеспечивает бит-точную воспроизводимость. Контрольный тест воспроизводимости подтвердил совпадение результатов до десятых долей метра при повторных прогонах одной конфигурации.

5.5 Главные результаты

Сводные результаты абляции на трёх наборах данных приведены в табл. 2. Формат данных в ячейках соответствует режимам «stdQ / modQ» – слева стандартная ковариация процесса, справа модифицированная. Все значения вычислены на тестовом сегменте каждого трека (последние 30 % кадров), не использованном при обучении EDL-модели, что исключает утечку обучающих данных в оценку траекторной точности. На всех трёх наборах лучшая EDL-конфигурация превосходит лучшую эвристическую базовую линию по средней траекторной ошибке, причём величина выигрыша монотонно возрастает с информативностью визуальной сцены.

На наборе D1 (высокая информативность) лучшая конфигурация *theoretical_clipped* обеспечивает Mean ATE 4,79 м против 8,82 м у эвристики – снижение на 46 %. Доля точных кадров (ошибка не более 5 м) возрастает с 27,1 % до 65,8 % (в 2,4 раза), доля кадров с ошибкой не более 10 м – с 66,6 % до 85,1 %.

Табл. 2. Сводные результаты абляционного исследования (тестовый сегмент, без утечки обучающих данных).

Table 2. Summary results of the ablation study (test segment, leakage-free).

Конфигурация	Mean ATE, м	Median, м	p95, м	<5 м, %	<10 м, %
Набор данных D1					
heuristic	8,82 / 11,66	8,29 / 10,72	19,91 / 20,87	27,1 / 14,0	66,6 / 47,1
rayleigh	8,80 / 7,48	8,12 / 5,90	18,42 / 16,24	26,0 / 34,4	65,3 / 78,4
theoretical	9,59 / 5,85	9,91 / 4,58	18,34 / 14,90	32,5 / 56,1	50,2 / 83,9
theoretical_clipped	4,80 / 4,79	2,96 / 2,91	13,66 / 13,64	65,8 / 65,8	85,1 / 85,1
Набор данных D2					
heuristic	16,73 / 18,00	16,50 / 18,52	28,75 / 28,60	5,5 / 2,5	19,0 / 16,4
rayleigh	17,97 / 16,19	18,48 / 16,72	29,12 / 26,63	3,2 / 2,9	14,4 / 22,8
theoretical	29,73 / 23,55	28,14 / 24,29	46,43 / 39,67	0,6 / 6,0	1,3 / 15,3
theoretical_clipped	17,72 / 15,50	17,96 / 15,79	29,17 / 27,37	3,8 / 5,6	16,2 / 25,9
Набор данных D3					
heuristic	3,75 / 3,75	1,49 / 1,49	12,54 / 12,54	72,5 / 72,5	86,7 / 86,7
rayleigh	3,66 / 3,66	1,63 / 1,63	12,08 / 12,08	74,2 / 74,2	88,5 / 88,5
theoretical	3,00 / 3,00	1,75 / 1,75	8,65 / 8,65	81,0 / 81,0	98,1 / 98,1
theoretical_clipped	3,33 / 3,33	1,92 / 1,92	9,59 / 9,59	79,2 / 79,2	96,0 / 96,0

На наборе D3 (средняя информативность, низковысотный полёт с манёврами) лучшая конфигурация *theoretical* обеспечивает Mean ATE 3,00 м против 3,75 м у эвристики – снижение на 20%. Преимущество наиболее выражено в хвосте распределения: p95 ATE снижается с 12,54 м до 8,65 м (на 31%), доля кадров с ошибкой не более 10 м растёт с 86,7% до 98,1%.

На наборе D2 (пониженная информативность, высотный полёт практически без манёвров) лучшая конфигурация *theoretical_clipped* обеспечивает Mean ATE 15,50 м против 16,73 м у эвристики – снижение на 7 %, при росте доли кадров с ошибкой не более 10 м с 19,0% до 25,9%. Прямое использование полной ковариации (формула *theoretical*) на этом наборе приводит к деградации (Mean ATE до 29,7 м), что обсуждается в разделе 5.6.

5.6 Выбор формулы адаптивной ковариации

Оптимальная формулировка адаптивной ковариации различается между наборами данных и определяется тем, насколько калибрована эпистемическая компонента NIG-выхода. На наборе D3, где эпистемическая и алеаторная компоненты сбалансированы, и полная ковариация калибрована (доля попаданий в один интервал составляет 76% при идеальных 68%), наилучший результат даёт формула *theoretical*. На наборах D1 и D2 эпистемическая компонента переоценена: она превосходит алеаторную в 6–17 раз, что является известным свойством эвиденциальной регрессии и не гарантирует строгой калибровки. В этом режиме прямое использование полной ковариации приводит к деградации траекторной точности; формула *theoretical_clipped*, привязывающая оценку к алеаторной компоненте и эвристическому приору, обеспечивает лучший результат и выступает необходимым механизмом робастности.

5.7 Верификация теоремы о вырождении наблюдаемости

Эффект модифицированной ковариации процесса на коэффициент усиления Калмана подтверждает теорему о вырождении наблюдаемости, но проявляется не единообразно, а в зависимости от соотношения частоты абсолютных обновлений и темпа накопления дрейфа. На наборе D3, где между абсолютными обновлениями накапливается заметный позиционный дрейф, модификация повышает эффективный коэффициент усиления с 0,85 до 0,97, что напрямую снижает траекторную ошибку. На наборе D1, где абсолютные обновления точны и достаточно часты, эффект на лучшую конфигурацию мал. Это уточняет область действия теоремы: модификация критична при низкочастотных абсолютных обновлениях и высоким дрейфе и второстепенна при частых точных коррекциях.

5.8 Границы применимости

Величина выигрыша эвиденциальной ковариации определяется тем, насколько контекстный вектор информативен относительно фактической ошибки локализации. Количественной характеристикой служит корреляция предсказанной моделью ошибки с фактической, монотонно связанная с достигаемым выигрышем: на наборе D1 она составляет 0,96 (выигрыш 46 %), на наборе D3 – 0,56 (выигрыш 20 %), на наборе D2 – 0,16 (выигрыш 7 %). На наборе D2 корреляция предсказания EDL (0,16) оказывается ниже корреляции эвристического предиктора (0,25): основными факторами вариации ошибки являются динамические маневры (корреляция с угловой скоростью курса 0,62, с креном 0,47), однако полёт выполнялся на постоянной высоте практически без манёвров, в силу чего предсказание модели вырождается в почти постоянное, а эпистемическая компонента раздувается.

Несмотря на это, формула *theoretical_clipped* сохраняет небольшое преимущество над эвристикой (7 %) за счёт привязки к алеаторной компоненте, что задаёт нижнюю границу применимости: при корреляции предсказания ниже корреляции эвристики выигрыш минимален, но не отрицателен. Верхняя граница достигается на наборе D1, где качество сопоставления существенно варьируется и модель точно различает надёжные и ненадёжные измерения. Таким образом, величина выигрыша предсказуемо определяется наблюдаемой вариацией факторов ошибки локализации, что согласуется с введённым понятием объяснимой наблюдаемости: метод эффективен тогда, когда контекст измерения действительно информативен о его точности.

6. Обсуждение

Полученные результаты показывают, что адаптивное формирование ковариации измерений на основе эвиденциального обучения даёт устойчивый выигрыш над эвристической моделью на всех протестированных наборах данных, однако величина этого выигрыша не является постоянной, а закономерно определяется информативностью контекста измерения. Ключевым предиктором оказалась корреляция предсказанной моделью ошибки с фактической: чем сильнее признаки сопоставления связаны с реальной точностью локализации, тем больший выигрыш обеспечивает адаптивная ковариация. Это придаёт методу свойство предсказуемости – по характеристикам полётных данных можно заранее оценить ожидаемую эффективность подхода.

Существенным практическим выводом является зависимость оптимальной формулировки ковариации от калибровки эпистемической компоненты. На доменах, где нормально-обратное гамма-распределение даёт калиброванную полную ковариацию, прямое её использование оптимально; на доменах с переоценённой эпистемической компонентой необходима регуляризация через привязку к алеаторной составляющей и эвристическому приору. Это согласуется с известными в литературе ограничениями эвиденциальной регрессии и подчёркивает, что робастная формулировка с зажимом является не

вспомогательным приёмом, а необходимым элементом метода при работе в неконтролируемых условиях.

Среди ограничений настоящего исследования следует отметить использование синтетических и полунатурных наборов данных, обучение и тестирование в пределах одного домена для каждой модели, а также оценку на ограниченном числе типов сцен. Эпистемическая компонента, несмотря на корректное ранжирование надёжности измерений, демонстрирует систематическую переоценку масштаба неопределённости на части доменов, что требует дополнительной калибровки.

7. Заключение

В работе предложен подход к адаптивному формированию ковариации измерений нейросетевого модуля абсолютной визуальной локализации в составе гибридной системы визуально-инерциальной навигации беспилотных летательных аппаратов. Подход основан на эвиденциальном глубоком обучении с параметризацией нормально-обратным гамма-распределением и позволяет в едином проходе оценивать алеаторную и эпистемическую компоненты неопределённости локализации. Введено понятие степени объяснимой наблюдаемости, доказаны теоремы об информационном выигрыше адаптивной ковариации, о вырождении наблюдаемости при стандартной ковариации процесса и о согласованности фильтра.

Систематическое экспериментальное исследование на двух синтетических и одном полётном наборах данных показало, что предложенный метод превосходит эвристическую базовую линию по средней траекторной ошибке на всех наборах, обеспечивая снижение от 7 до 46 % в зависимости от информативности сцены. Установлен количественный критерий применимости: величина выигрыша монотонно связана с корреляцией предсказанной ошибки с фактической. Показано, что формула с зажимом обеспечивает робастность при нарушении калибровки эпистемической компоненты, а модификация ковариации процесса повышает коэффициент усиления Калмана в режимах с накоплением дрейфа между обновлениями.

Перспективными направлениями развития работы являются междоменное обучение эвиденциальной модели с явным учётом доменного разрыва, пост-калибровка эпистемической компоненты для восстановления корректного масштаба неопределённости, расширение контекстного вектора признаками динамики полёта, расширенная валидация на натурных полётных данных и аппаратная реализация модуля на встраиваемых вычислителях с оценкой работы в реальном времени.

Список литературы / References

- [1] Arafat M.Y., Alam M.M., Moh S. Vision-based navigation techniques for unmanned aerial vehicles: review and challenges. *Drones*, 2023, vol. 7, no. 2, article 89, pp. 1-33. DOI: 10.3390/drones7020089.
- [2] Couturier A., Akhloufi M.A. A review on absolute visual localization for UAV. *Robotics and Autonomous Systems*, 2021, vol. 135, article 103666, pp. 1-28. DOI: 10.1016/j.robot.2020.103666.
- [3] Mourikis A.I., Roumeliotis S.I. A multi-state constraint Kalman filter for vision-aided inertial navigation. In Proc. of the IEEE International Conference on Robotics and Automation (ICRA), 2007, pp. 3565-3572. DOI: 10.1109/ROBOT.2007.364024.
- [4] Potje G., Cadar F., Araujo A., Martins R., Nascimento E.R. XFeat: accelerated features for lightweight image matching. In Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024, pp. 2682-2691. DOI: 10.1109/CVPR52733.2024.00259.
- [5] Lindenberger P., Sarlin P.-E., Pollefeys M. LightGlue: local feature matching at light speed. In Proc. of the IEEE/CVF International Conference on Computer Vision (ICCV), 2023, pp. 17627-17638. DOI: 10.1109/ICCV51070.2023.01616.

- [6] Geneva P., Eckenhooff K., Lee W., Yang Y., Huang G. OpenVINS: a research platform for visual-inertial estimation. In Proc. of the IEEE International Conference on Robotics and Automation (ICRA), 2020, pp. 4666-4672. DOI: 10.1109/ICRA40945.2020.9196524.
- [7] Qiu Y., Chen Y., Zhang Z., Wang W., Scherer S. MAC-VO: metrics-aware covariance for learning-based stereo visual odometry. In Proc. of the IEEE International Conference on Robotics and Automation (ICRA), 2025, pp. 3803-3814. DOI: 10.1109/ICRA55743.2025.11128482.
- [8] Blundell C., Cornebise J., Kavukcuoglu K., Wierstra D. Weight uncertainty in neural networks. In Proc. of the 32nd International Conference on Machine Learning (ICML), 2015, pp. 1613-1622.
- [9] Gal Y., Ghahramani Z. Dropout as a Bayesian approximation: representing model uncertainty in deep learning. In Proc. of the 33rd International Conference on Machine Learning (ICML), 2016, pp. 1050-1059.
- [10] Levy A., Klein I. Adaptive neural unscented Kalman filter. *IEEE Transactions on Intelligent Vehicles*, 2026, vol. 11, no. 3, pp. 428-437. DOI: 10.1109/TIV.2026.3657014.
- [11] Versano G., Klein I. A hybrid neural-assisted unscented Kalman filter for unmanned ground vehicle navigation. arXiv preprint, 2026. DOI: 10.48550/arXiv.2603.11649.
- [12] Amini A., Schwarting W., Soleimany A., Rus D. Deep evidential regression. *Advances in Neural Information Processing Systems (NeurIPS)*, 2020, vol. 33, pp. 14927-14937.
- [13] Han J., Beyer L.L., Cavalheiro G.V., Karaman S. SCREP: Scene Coordinate Regression and Evidential Learning-based Perception-Aware Trajectory Generation. arXiv preprint, 2025. DOI: 10.48550/arXiv.2507.07467.
- [14] Qin T., Li P., Shen S. VINS-Mono: a robust and versatile monocular visual-inertial state estimator. *IEEE Transactions on Robotics*, 2018, vol. 34, no. 4, pp. 1004-1020. DOI: 10.1109/TRO.2018.2853729.
- [15] Leutenegger S., Lynen S., Bosse M., Siegwart R., Furgale P. Keyframe-based visual-inertial odometry using nonlinear optimization. *The International Journal of Robotics Research*, 2015, vol. 34, no. 3, pp. 314-334. DOI: 10.1177/0278364914554813.
- [16] Bloesch M., Burri M., Omari S., Hutter M., Siegwart R. Iterated extended Kalman filter based visual-inertial odometry using direct photometric feedback. *The International Journal of Robotics Research*, 2017, vol. 36, no. 10, pp. 1053-1072. DOI: 10.1177/0278364917728574.
- [17] Sarlin P.-E., DeTone D., Malisiewicz T., Rabinovich A. SuperGlue: learning feature matching with graph neural networks. In Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 4938-4947. DOI: 10.1109/CVPR42600.2020.00499.
- [18] Shi J., Cheng J., Li Q., Lu X., Zhang C., Li L. Towards GNSS-Free UAV Localization via XFeatLightGlue and Semantic Embedding. In Proc. of the IEEE 6th International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE), 2025, pp. 192-197. DOI: 10.1109/ICBASE66587.2025.11181328.
- [19] Akhoury S. Real-time localization of planar targets on power-constrained platforms. M.Sc. Thesis, University of Ottawa, 2017, 153 p. Available at: https://www.site.uottawa.ca/~laganier/publications/thesis/Sharat_Akhoury_Thesis.pdf, accessed 24.05.2026.
- [20] Mohamed A.H., Schwarz K.P. Adaptive Kalman filtering for INS/GPS. *Journal of Geodesy*, 1999, vol. 73, no. 4, pp. 193-203. DOI: 10.1007/s001900050236.
- [21] Dahan Y., Revach G., Dunik J., Shlezinger N. Bayesian KalmanNet: quantifying uncertainty in deep learning augmented Kalman filter. *IEEE Transactions on Signal Processing*, 2025, vol. 73, pp. 2558-2572. DOI: 10.1109/TSP.2025.3581703.
- [22] Wu W., Zhang C., Woodland P.C. Estimating the uncertainty in emotion attributes using deep evidential regression. In Proc. of the 61st Annual Meeting of the Association for Computational Linguistics (ACL), Volume 1: Long Papers, 2023, pp. 15681-15695. DOI: 10.18653/v1/2023.acl-long.873.
- [23] Chinta A., Tran L.V., Katukuri J. ProbFM: Probabilistic Time Series Foundation Model with Uncertainty Decomposition. arXiv preprint, 2026. DOI: 10.48550/arXiv.2601.10591.
- [24] Ben Ayed S., Sadeghian Broujey R., Hamza R. Remaining useful life prediction with uncertainty quantification using evidential deep learning. *Journal of Artificial Intelligence and Soft Computing Research*, 2025, vol. 15, no. 1, pp. 37-55. DOI: 10.2478/jaiscr-2025-0003.
- [25] Pasqualetto Cassinis L., Park T.H., Stacey N., D'Amico S., Menicucci A., Ahms I., Sanchez-Gestido M. Leveraging neural network uncertainty in adaptive unscented Kalman filter for spacecraft pose estimation. *Advances in Space Research*, 2023, vol. 71, no. 12, pp. 5061-5082. DOI: 10.1016/j.asr.2023.02.021.

- [26]. Bar-Shalom Y., Li X.R., Kirubarajan T. Estimation with Applications to Tracking and Navigation: Theory, Algorithms and Software. New York, John Wiley & Sons, 2001, 584 p. DOI: 10.1002/0471221279.ch11.
- [27]. Li M., Zhang Y. Deep learning-assisted ES-EKF for surface AUV navigation with photometric feedback. *Journal of Marine Science and Engineering*, 2025, vol. 13, no. 11, article 2035. DOI: 10.3390/jmse13112035.
- [28]. Sola J. Quaternion kinematics for the error-state Kalman filter. arXiv preprint arXiv:1711.02508, 2017, 93 p. Available at: <https://arxiv.org/abs/1711.02508>, accessed 24.05.2026.

Информация об авторах / Information about authors

Полина Александровна ЛАЗАРЕВА – кандидат физико-математических наук, доцент кафедры автоматизации и управления, ведущий научный сотрудник научно-исследовательской лаборатории № 27 КАИ. Сфера научных интересов: компьютерное зрение, автономная навигация беспилотных летательных аппаратов, робототехника, машинное обучение.

Polina Aleksandrovna LAZAREVA – Cand. Sci. (Phys.-Math.), associate professor at the Department of Automation and Control, and leading researcher at Research Laboratory No. 27 of the Kazan National Research Technical University named after A.N. Tupolev – KAI. Her research interests include computer vision, autonomous navigation of unmanned aerial vehicles, robotics, and machine learning.

Николай Витальевич ГУРЬКО – инженер кафедры автоматизации и управления, лаборант научно-исследовательской лаборатории № 27 КАИ. Сфера научных интересов: компьютерное зрение, автономная навигация беспилотных летательных аппаратов, робототехника, машинное обучение. Сфера научных интересов: компьютерное зрение, автономная навигация беспилотных летательных аппаратов, робототехника, машинное обучение, распознавание образов.

Nikolay Vitalyevich GURKO is an engineer in the Department of Automation and Control and a laboratory assistant at Research Laboratory No. 27 at the Kazan National Research Technical University named after A.N. Tupolev – KAI. His research interests include computer vision, autonomous navigation of unmanned aerial vehicles, robotics, machine learning, and pattern recognition.