УДК 681.5
EDN LIQJHA

А. АССАД, С. А. СЕРИКОВ

АДАПТАЦИЯ КОВАРИАЦИИ ШУМА В ОБОБЩЕННОМ ФИЛЬТРЕ КАЛМАНА С ИСПОЛЬЗОВАНИЕМ ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ ДЛЯ БОЛЕЕ ТОЧНОГО ОПРЕДЕЛЕНИЯ УГЛОВ ОРИЕНТАЦИИ БПЛА

Точное определение углов ориентации беспилотных летательных аппаратов (БПЛА) имеет решающее значение для автономной навигации, особенно при использовании лишь измерений гироскопов, акселерометров и магнитометров без привлечения данных глобальной системы позиционирования (GPS). Перспективным представляется метод обучения искусственного интеллекта с подкреплением (ОП), который позволяет повысить эффективность применяемого при определении углов ориентации обобщенного фильтра Калмана (ОФК). Предлагаемый подход предусматривает привлечение модели Q-обучения и стратегии поиска наилучшего решения для коррекции матрицы ковариации шума измерений в ОФК в автономном режиме. За счет механизма вознаграждения, стимулирующего действия, с помощью которых сводится к минимуму погрешность прогнозирования углов ориентации относительно истинных измерений, ОП дает возможность динамически оптимизировать матрицу ковариации шума измерений. Интегрированный алгоритм ОП и ОФК (далее – ОП-ОФК) был реализован и протестирован. Результаты показывают, что он значительно превосходит традиционный ОФК в части определения углов ориентации БПЛА.

Ключевые слова: беспилотный летательный аппарат, обучение с подкреплением, обобщенный фильтр Калмана, оценивание ориентации.

Введение

Обучение с подкреплением (ОП) – разновидность машинного обучения, в котором агент учится находить наилучшую стратегию для достижения своих целей, взаимодействуя с окружающей средой и получая при этом вознаграждение за наиболее удачные решения. Существуют три вида машинного обучения: контролируемое, неконтролируемое и ОП. К последнему относятся такие понятия, как обучение с использованием функции полезности (*Q*-обучение), глубокое ОП и глубокая *Q*-сеть, то есть нейросеть, обученная методом глубокого ОП. *Q*-обучение – это безмодельный алгоритм, базирующийся на функции ценности, нестрого регламентированный и задействуемый для поиска оптимальной стратегии действий агента в заданных

Ассад Аммар. Аспирант, Санкт-Петербургский государственный университет аэрокосмического приборостроения.
Сериков Сергей Анатольевич. Доктор технических наук, доцент, Санкт-Петербургский государственный университет аэрокосмического приборостроения.

условиях. Алгоритм определяет наилучшую последовательность действий, которые агент должен предпринять с учетом своего текущего состояния. Латинская буква Q в названиях понятий, связанных с ОП, обозначает «качество» (quality) и указывает на то, что от правильно выбранного действия зависит будущее вознаграждение.

Q-обучение широко применяется в различных областях, таких как беспилотные технологии, робототехника, игры и т.д. Во многих случаях ОП может выступать в качестве как дополнительной функции другого алгоритма, так и самостоятельного метода [1, 2]. В последнее время была найдена возможность использования ОП в неопределенных условиях – при наличии шумов, помех, дрейфа, смещения нуля датчика, неблагоприятных факторов окружающей среды (ветер, температурные перепады, магнитные возмущения), временных задержек, ошибок интегрирования и т.д. Беспилотные летательные аппараты (БПЛА) функционируют именно в таких неопределенных условиях, поэтому точное определение углов их ориентации имеет решающее значение для автономной навигации, чтобы обеспечить надежное управление их движением. С этой целью чаще всего задействуется фильтр Калмана (ФК) [3, 4].

Следует отметить, что ФК по-прежнему нуждается в усовершенствовании, поскольку погрешности измерительных датчиков влияют на результаты оценивания параметров. Кроме того, полет БПЛА может быть стабильным или сопровождаться маневрированием, поэтому акселерометры регистрируют различное кажущееся ускорение (общее линейное ускорение за вычетом ускорения силы тяжести). В случае совершения БПЛА резких маневров акселерометр не только измеряет линейное ускорение, но и испытывает воздействие других сил, например центробежных или кориолисовых. Если использовать эти данные в ФК с постоянной матрицей R (матрица ковариации шума измерений), возрастает вероятность некорректной оценки углов ориентации. Постоянная матрица R не учитывает изменения в шуме измерений, обусловленные воздействием дополнительных факторов, что приводит к неточности определения углов ориентации БПЛА. Чтобы этого избежать, необходима адаптация матрицы R, которая осуществляется с помощью ОП.

Для оценивания углов ориентации БПЛА требуются входные данные гироскопов, акселерометров и магнитометров, которые затем обрабатываются в ФК для получения параметров ориентации в виде углов Эйлера или кватернионов. Обычно при использовании модели системы процедура калмановской фильтрации состоит из двух этапов: на первом выполняется априорное прогнозирование (или прогноз вектора состояния), на втором – его коррекция (обновления) с помощью модели измерений. В обобщенном ФК (ОФК) учитывается модель линеаризованной системы и модель измерений относительно оцениваемого вектора состояния. Матрица ковариации шума измерений (обозначаемая как R и подлежащая адаптации) и матрица состояния или ковариации шума процесса (обозначаемая в данной работе как Q_f вместо Q во избежание путаницы с термином «Q-обучение») задействуются на обоих этапах ОФК [5, 6].

Как правило, при вычислении углов ориентации методом калмановской фильтрации матрицы R и Q_f полагаются постоянными, что справедливо лишь в режиме установившегося полета, который не может быть гарантирован в реальном времени из-за наличия шумов, помех, погрешностей измерений и т.д. На ковариацию шума измерений влияют период времени, место и маневрирование. Таким образом, даже если истинная матрица ковариации известна, иногда (например, при маневрировании) измерения полезно не учитывать в калмановских расчетах, поскольку в этих слу-

чаях модель измерений недействительна. Предложенный в настоящей работе метод адаптации матриц ковариации шумов датчиков на этапе измерений предусматривает использование ОП на основе *Q*-обучения. Это позволяет в каждый момент времени найти адекватное ситуации значение матрицы *R,* динамически меняющейся в процессе полета. Статья имеет следующую структуру: в первом разделе представлен обзор современного состояния соответствующей области исследований; во втором подробно рассматривается задача оценивания углов ориентации БПЛА с помощью ОФК. Третий раздел посвящен методу *Q*-обучения; в четвертом приводятся результаты его применения на практике; в заключении представлены основные выводы.

1. Современное состояние

В работе [7] предложен алгоритм оценивания вектора состояния системы, сочетающий ОФК и *Q*-обучение, в котором стратегия адаптации ковариации базируется на системе вознаграждений, или «*Q*-ценностей». Метод получил название «ОФК с *Q*-обучением» (QL-EKF) и, по мнению авторов, оказался менее чувствительным к неопределенности ковариации шума. Для подтверждения его эффективности было проведено численное моделирование автономной навигации космического аппарата, результаты которого показали, что новый алгоритм работает лучше традиционного ОФК.

Авторы [8] предложили метод на основе *Q*-обучения для автоматической настройки матрицы ковариации шумов измерений и возмущений. Формируется механизм ОП, который управляет заранее определенным набором возможных матриц ковариации шума таким образом, чтобы в него можно было включить пару матриц с наименьшей разницей между выходными и прогнозными значениями измерений. Эффективность *Q*-обучения подтверждается на реальных полетных данных, полученных с БПЛА и обработанных методом Монте-Карло, который применяется для оценивания углов ориентации с помощью ОФК.

В [9] углы ориентации рассчитывались с помощью интегрированного алгоритма – левоинвариантного ОФК (LI-EKF) в сочетании с адаптивным алгоритмом оценивания ковариации шума. Последний представляет собой итеративный алгоритм максимизации ожидания для адаптации изменяющихся во времени параметров шума. В [10] рассмотрена система нечеткого вывода для адаптации матрицы ковариации шума измерений в ФК с помощью акселерометров и гироскопов. Разработанный метод оказался лучше обычного ФК. С той же целью система нечеткого вывода была применена и в [11], при этом входными данными служили измерения гироскопов, акселерометров и магнитометров. Углы ориентации пикоспутника вычислялись с помощью ансцентного ФК (АФК), недостатком которого является снижение точности и расходимость с течением времени в случае системной неопределенности или некорректного выполнения измерений. В связи с этим в работе [12] был построен отказоустойчивый алгоритм оценивания углов ориентации на основе робастного адаптивного АФК, способный корректировать ковариацию порождающего шума (*Q*-адаптация) или шума измерений (*R*-адаптация) в зависимости от типа неисправности. Новая адаптационная схема, разработанная для традиционного АФК, позволяет обнаружить и изолировать неисправность, после чего выполняется обязательная адаптация в соответствии с типом дефекта.

В исследовании [5], чтобы скорректировать матрицу ковариации шума измерений, в качестве входных данных для системы нечеткого ввода использовались показания акселерометров и магнитометров. Недостаток этого подхода состоит в том, что адаптивный метод требует выполнения измерений, в то время как при подходе на основе Q-обучения процесс адаптации начинается с произвольного значения матрицы ковариации шума измерений, которое затем редактируется до достижения оптимальной величины.

В [13] предложен улучшенный алгоритм IMM-KF (Interacting Multiple Models-Kalman Filter – интерактивный многомодельный ФК), который учитывает априорную информацию о планируемом маршруте дрона и данные о его координатах, скорости и ускорении и позволяет прогнозировать его местоположение в реальном времени. Наличие нескольких интерактивных моделей, в частности, помогает сформировать конфигурацию фильтра, чтобы установить параметры и выбрать для каждого из них оптимальное значение.

В настоящей работе предлагается подход, сочетающий Q-обучение и ОФК, для более точного определения углов ориентации БПЛА с использованием данных инерциального измерительного модуля в отсутствие сигналов GPS. В принципе, для решения этой задачи достаточно только ОФК, однако Q-обучение позволяет снизить погрешность оценивания углов и обеспечивает динамическую адаптацию матрицы ковариации измерений R к внешним условиям. Предполагается, что с помощью нового метода можно существенно повысить точность определения параметров ориентации в сравнении с традиционными подходами на основе ОФК, эффективность которых оценивается по стандартным показателям, например по среднему значению погрешности или по среднеквадратической погрешности.

2. Оценивание углов ориентации

Постановка задачи

Введем вектор состояния с четырьмя переменными (он описывает ориентацию БПЛА), который содержит кватернион $\overline{q}$ поворота между системой координат n (север–восток–вниз) и связанной с БПЛА системой координат b. При этом для удобства углы ориентации БПЛА обычно представлены в виде углов Эйлера или кватернионов; линейность ориентации БПЛА объясняется ниже. Углы Эйлера более удобны при разложении поворотов объекта на отдельные степени свободы (например, для кинематических узловых соединений), однако у них есть недостатки – неоднозначность и проблема «складывания рамок». На практике предпочтительнее использовать кватернионы, поскольку они эффективнее и доступнее для машинных вычислений.

При работе с углами Эйлера приходится перемножать их значения для трех последовательно выполненных поворотов, тогда как в ситуации с кватернионом мы имеем дело только с одним поворотом, а поскольку он уже содержит функции синуса и косинуса, то преобразовать кватернион в матрицу достаточно просто. Любая последовательность поворотов может быть представлена как непрерывная траектория кватерниона без каких-либо нарушений целостности, что имеет место в случае углов Эйлера. Кроме того, в отличие от углов Эйлера кватернионы требуют лишь простых операций с числами, а не сложных вычислений синусов и косинусов. К тому же кватернионы обладают двойственностью [14], т.е. могут описывать одну

и ту же ориентацию двумя различными способами. Предположим, имеется вектор состояния x_k:

$$x_k = q_k = \begin{bmatrix} q_{1,k} & q_{2,k} & q_{3,k} & q_{4,k} \end{bmatrix}^T,$$

где q_4 – скалярная, а $[q_1\ q_2\ q_3]$ – векторная части, вычисляемые по показаниям гироскопов, при этом ω_i – истинные измерения гироскопа, а $\eta_{g,i}$ – шум (по осям $i = x, y, z$). Тогда данные гироскопа $\tilde{\omega}_i$ будут иметь вид:

$$\tilde{\omega}_i = \omega_i + \eta_{g,i},\ i = x, y, z.$$

При оценивании углов ориентации можно пренебречь угловой скоростью вращения Земли (~7,2921 × 10^{-5} рад/с), поскольку она пренебрежительно мала по сравнению с величинами, обычно измеряемыми гироскопами во время полета БПЛА. Будем полагать, что шумы в измерениях гироскопов, акселерометров и магнитометров – невзаимосвязанные центрированные некоррелированные гауссовские белые шумы. Обозначим измеренные значения ускорений как $\tilde{f}_i$, истинные – как f_i, а шум – как $\eta_{a,i}$ (по осям $i = x, y, z$):

$$\tilde{f}_i = f_i + \eta_{a,i},\ i = x, y, z.$$

Аналогичным образом измеренные значения магнитного поля – $\tilde{M}_i$, истинные – M_i, а шум – $\eta_{m,i}$ (по осям $i = x, y, z$):

$$\tilde{M}_i = M_i + \eta_{m,i},\ i = x, y, z.$$

Вектор измерений $\tilde{z}_k$ состоит из данных акселерометров $f_{x,y,z,k}$ и магнитометров $M_{x,y,z,k}$:

$$\tilde{z}_k = \begin{bmatrix} \tilde{f}_{xyz,k} \\ \tilde{M}_{xyz,k} \end{bmatrix} = \begin{pmatrix} \tilde{f}_{x,k} & \tilde{f}_{y,k} & \tilde{f}_{z,k} & \tilde{M}_{x,k} & \tilde{M}_{y,k} & \tilde{M}_{z,k} \end{pmatrix}^T. \quad (1)$$

Величины $\tilde{z}_k$ и $\tilde{\omega}$ используются на разных этапах (см. ниже) оценивания вектора состояния с помощью ОФК, а затем и Q-обучения, чтобы обеспечить получение надежных результатов при наличии шума. На практике все инерциальные и магнитные датчики, в том числе гироскопы, акселерометры и магнитометры, подвержены различным воздействиям, в том числе перепадам температуры, механическим вибрациям, электромагнитным явлениям (как внешним, так и связанным с несущей платформой). На них влияют также факторы, относящиеся непосредственно к самим датчикам, – нестабильность смещения нуля и погрешности масштабного коэффициента. Поскольку алгоритм ОФК учитывает шумы в моделях процесса и измерений, такие неопределенности неявно включаются в уравнения фильтра, что будет рассмотрено в следующем разделе.

ОФК для оценивания углов ориентации

Динамическая оценка углов ориентации движущегося объекта выполняется различными системами и может быть представлена в линейном виде. Однако в случае БПЛА возникает ряд сложностей, обусловленных высокой маневренностью аппаратов, быстрым изменением их угловой скорости и необходимостью производить точные вычисления под воздействием внешних факторов, например ветра. Наличие этих проблем

требует применения усовершенствованных методов фильтрации, таких как ФК, которые позволяют учитывать нелинейные преобразования в динамике вращения. Предположим, вектор состояния x содержит кватернион, тогда уравнения состояния и измерений в момент времени t_k можно записать соответственно как (2) и (4) [5] (табл. 1).

Таблица 1

Уравнения моделей состояния и измерений для оценивания ориентации БПЛА

Динамика системы	$x_k = f(x_{k-1}, W_k),$	(2)
	$W_k \sim N(0, Q_{fk})$	(3)
Модель измерений	$\tilde{z}_k = h(x_k, V_k),$	(4)
	$V_k \sim N(0, R_k)$	(5)

Уравнение (2) представляет собой уравнение состояния системы, $f(\cdot)$ – нелинейная векторная функция, x_k – вектор состояния $n \times 1$ в момент времени t_k, W_{k-1} – шум возмущений $n \times 1$ с неизвестной диагональной матрицей ковариации Q_{fk} (с заданной дисперсией σ_g^2); (4) – уравнение измерений, $h(\cdot)$ – нелинейная векторная функция, $\tilde{z}_k$ – вектор измерений $m \times 1$ в момент времени t_k, V_k – аддитивный шум измерений $m \times 1$ с неизвестной диагональной матрицей ковариации R_k (с заданной дисперсией σ_a^2 для акселерометров и σ_m^2 для магнитометров); $\{x_0, W_k, V_k\}$ считаются некоррелированными случайными величинами.

Для ОФК: уравнения (2) и (4) описывают нелинейную систему, где ФК применяется различными способами. Одним из самых простых путей является линеаризация динамических параметров системы и функции измерений в соответствии с оценкой состояния $\hat{x}$ [15]. Такой подход и называется обобщенной калмановской фильтрацией. Параметр x_k – истинное состояние, $\hat{x}_k^-$ – априорная оценка x_k (результат фазы прогноза), $\hat{x}_k^+$ – апостериорная оценка x_k ($\hat{x}_k^+ = \hat{x}_k$). Уравнения измерений и динамики системы для ОФК представлены в табл. 2.

Таблица 2

Линеаризованные модели системы и измерений

Динамика системы	$x_k = F_k x_{k-1} + G_k W_k,$	(6)
	$F_k = \left.\frac{\partial f}{\partial x}\right\|_{x=x_{k-1}}, \qquad G_k = \left.\frac{\partial f}{\partial W}\right\|_{x=x_{k-1}},$	(7)
	F_k – якобиан модели состояния (матрица переходного состояния), G_k – входная матрица шума возмущений. Модель линеаризации находится вокруг точки x_{k-1}	
Модель измерений	$\tilde{z}_k \approx H_k x_k + V_k,$	(8)
	$H_k = \left.\frac{\partial h}{\partial x}\right\|_{x=\hat{x}_k^-},$	(9)
	H_k – якобиан модели измерений	

В табл. 3 приведены уравнения Калмана.

Таблица 3

Основные этапы калмановской фильтрации

Прогноз	$\hat{x}_k^- = f(\hat{x}_{k-1}^+)$	
Априорная ковариация погрешности оценки состояния	$P_k^- = F_{k-1}P_{k-1}^+F_{k-1}^T + G_{k-1}Q_{f,k-1}G_{k-1}^T$, P_k^- – априорная ковариация погрешности оценки состояния в выборке k, P_{k-1}^+ – апостериорная ковариация погрешности оценки состояния в выборке k–1	(10)
Коэффициент усиления Калмана	$K_k = P_k^- H_k^T \left(H_k P_k^- H_k^T + R_k\right)^{-1}$, K_k – коэффициент усиления Калмана в выборке k	(11)
Погрешность (невязка) измерений	$\delta z_k^- = \tilde{z}_k - \hat{z}_k$, $\hat{z}_k = H_k \hat{x}_k^-$	(12)
Обновление состояния (стадия коррекции)	$\hat{x}_k^+ = \hat{x}_k^- + K_k \cdot \delta z_k^-$	(13)
Обновление апостериорной ковариации состояния	$P_k^+ = \left(I_{n\times n} - K_k H_k\right) \cdot P_k^- \cdot \left(I_{n\times n} - K_k H_k\right)^T + K_k R_k K_k^T$, P_k^+ – апостериорная ковариация погрешности оценки состояния в выборке k	(14)

Параметр $\hat{z}_k$ – оценка измерений с помощью $\hat{x}_k^-$; $\tilde{z}_k$ – истинный вектор измерений, теоретически зависимый от истинного вектора состояния x. Существуют и другие формулы для уточнения матрицы P_k^+, но выбрана именно эта, поскольку содержит информацию матрицы R [6]. Добавление члена KRK^T позволяет сохранить положительный полуопределенный характер матрицы ковариации. Это свойство важно для корректной статистической интерпретации и обеспечения неотрицательности дисперсий. Уравнение оценивания априорного вектора состояния (кватерниона) имеет вид [16]:

$$\dot{q} = \frac{1}{2}\Omega(\tilde{\omega})q = \frac{1}{2}\begin{pmatrix} -[\tilde{\omega}\times] & \tilde{\omega} \\ -\tilde{\omega}^T & 0 \end{pmatrix} q \qquad [\tilde{\omega}\times] = \begin{pmatrix} 0 & -\tilde{\omega}_z & \tilde{\omega}_y \\ \tilde{\omega}_z & 0 & -\tilde{\omega}_x \\ -\tilde{\omega}_y & \tilde{\omega}_x & 0 \end{pmatrix}. \tag{15}$$

Используя вектор W_k в (2), перепишем уравнение (15):

$$\dot{q} = \frac{1}{2}\begin{pmatrix} 0 & \omega_z + \eta_{g,z} & -\omega_y - \eta_{g,y} & \omega_x + \eta_{g,x} \\ -\omega_z - \eta_{g,z} & 0 & \omega_x + \eta_{g,x} & \omega_y + \eta_{g,y} \\ \omega_y + \eta_{g,y} & -\omega_x - \eta_{g,x} & 0 & \omega_z + \eta_{g,z} \\ -\omega_x - \eta_{g,x} & -\omega_y - \eta_{g,y} & -\omega_z - \eta_{g,z} & 0 \end{pmatrix} q. \tag{15a}$$

Белый шум η_g отсутствует в уравнении состояния в качестве аддитивного шума, и уравнение состояния является линейным относительно кватерниона и нелинейным относительно шума. Чтобы преобразовать все измерения в систему координат n, рассчитаем матрицу прямых косинусов по четырем компонентам [17]:

$$C_n^b = \begin{pmatrix} 1-2q_2^2-2q_3^2 & 2q_1q_2+2q_3q_4 & 2q_1q_4-2q_2q_4 \\ 2q_1q_2-2q_3q_4 & 1-2q_1^2-2q_3^2 & 2q_1q_4+2q_2q_3 \\ 2q_1q_3+2q_2q_4 & 2q_1q_4-2q_2q_3 & 1-2q_1^2-2q_2^2 \end{pmatrix}^T. \tag{16}$$

Модель измерений имеет вид:

$$\begin{gathered} \tilde{z}_k = h(x) + V_k = \begin{bmatrix} C_n^b f_n \\ C_n^b M_n \end{bmatrix} + V_k, \\ V_k = \begin{bmatrix} \eta_{a,k} \\ \eta_{m,k} \end{bmatrix} = \begin{bmatrix} \eta_{ax,k} & \eta_{ay,k} & \eta_{az,k} & \eta_{mx,k} & \eta_{my,k} \end{bmatrix}^T, \end{gathered} \tag{17}$$

где M_n – магнитное поле в локальной системе координат, точно известное по всемирной модели магнитного поля Земли [19]; f_n – кажущееся ускорение в системе координат n, которое можно успешно аппроксимировать в ней вектором силы тяжести с противоположным знаком [0; 0; –g]. Эта аппроксимация не учитывает инерциальные ускорения, возникающие, например, при линейном движении объекта или под воздействием внешних помех, вклад таких ускорений учитывается в составе шума измерений V_k. Для рассматриваемого объекта ожидаемые инерциальные ускорения относительно невелики и составляют, как правило, порядка 0,01–0,1 g ($g = 9{,}81 \frac{м}{с^2}$). При такой малой величине упрощенное представление f_n остается востребованным для практических целей. Более подробные уравнения и иллюстрации приведены в [5]. При этом функции $f(\cdot)$ и $h(\cdot)$ можно записать следующим образом:

$$\begin{gathered} f(\cdot) = \frac{1}{2}\Omega\left(\tilde{\omega}(t)\right)x(t), \\ h(\cdot) = \begin{bmatrix} C_n^b f_n \\ C_n^b M_n \end{bmatrix}, \end{gathered} \tag{18}$$

C_n^b в выборке k рассчитывается для каждой выборки с использованием $x_k = [q_{1,k}\, q_{2,k}\, q_{3,k}\, q_{4,k}]$ (16). Независимо от того, в каком диапазоне – временном или дискретном – производятся вычисления [18], $\tilde{\omega}_k = \tilde{\omega}(t)$ является постоянной в течение периода интегрирования dt_k (номинально 16 мс):

$$F_k = \exp\left(\frac{dt_k}{2}\Omega\left(\tilde{\omega}_k\right)\right).$$

Матрицы Q_f и R представим как

$$Q_f = \sigma_g^2 . I_{4\times4}, \; R = \begin{bmatrix} \sigma_a^2 . I_{3\times3} & O_{3\times3} \\ O_{3\times3} & \sigma_m^2 . I_{3\times3} \end{bmatrix},$$

где $I_{a,b}$ – единичная матрица размерностью [a, b], а $O_{a,b}$ – нулевая матрица размерностью [a, b]. Значения σ_g, σ_a и σ_m берутся из технических описаний задействуемых датчиков. В разработанной методике σ_a и σ_m применяются при первой инициализа-

ции, а затем матрица R преобразуется в соответствии с принципами Q-обучения. Матрица G согласно (7) и посредством (15а) может быть выражена так:

$$G_k = \begin{pmatrix} q_{4,k} & -q_{3,k} & q_{2,k} \\ q_{3,k} & q_{4,k} & -q_{1,k} \\ -q_{2,k} & q_{1,k} & q_{4,k} \\ -q_{1,k} & -q_{2,k} & -q_{3,k} \end{pmatrix}.$$

Ранее предполагалось, что шум присутствует во всех измерениях датчиков. Хотя вся система в целом работает нелинейно, ОФК применялся с использованием локальной линеаризации вокруг номинальной точки x_k. Такое упрощение характерно для систем, где отклонение от номинальной траектории невелико и при этом большое значение имеет вычислительная эффективность. Несмотря на то что многие аспекты системы были упрощены или идеализированы, уровень шума в измерениях учитывался по-прежнему, поскольку вносит критический вклад в оценку состояния. В частности, влияние нелинейного шума модели измерений может существенно снизить точность работы ОФК.

3. Обучение с подкреплением и ОФК

Как уже было сказано, метод Q-обучения основан на принципе максимального вознаграждения в пространстве состояний, как правило состоящего из дискретных пар «состояние–действие», каждой из них присваивается скалярная величина, называемая Q-значением. Q-обучение позволяет адаптировать и скорректировать матрицу R, которая в стационарных условиях должна иметь постоянное значение (в соответствии с техническими характеристиками акселерометров и магнитометров). В ситуации с БПЛА такие условия не могут быть гарантированы, и под воздействием разных факторов, которыми сопровождается полет, R является переменной. В этом случае с помощью Q-обучения можно адаптировать матрицу для компенсации воздействия условий и среды полета.

Определение вознаграждения

Основная цель обучения с подкреплением – найти оптимальную стратегию, то есть такую плотность вероятности действий в каждом из состояний, которая обеспечивает максимальную отдачу при заданном коэффициенте ценности. Принцип Q-обучения зиждется на выборе правильных действий, за которые агент получит максимальное вознаграждение, и, наоборот, исключении неправильных, за которые награда не следует (т.е. следует наказание). При оценивании углов ориентации с помощью ОФК потенциальные вознаграждения могут быть следующими.

1. За действия, связанные с попыткой минимизировать вектор невязки измерений:

$$\text{Вознаграждение} = -\text{норма}\left(\delta z_k^-\right). \tag{19}$$

Здесь «норма» – это норма вектора, соответствующая формуле (допустим, вектор $L = [L_1, L_2, \ldots, L_{nv}]$), где nv – количество выборок, которое должно быть меньше, чем окно размерностью M (см. ниже):

$$\text{норма} = \left(\sum_{i=1}^{nv} L_i^2 \right)^{\frac{1}{2}}.$$

Это вознаграждение побуждает алгоритм сокращать разницу между прогнозируемыми и фактическими измерениями. От него напрямую зависит погрешность оценивания вектора состояния. В связи с этим оно чувствительно к шуму измерений, что устраняется с помощью ФК, регулирующего уровень шума для достижения баланса между шумом возмущений и шумом измерений.

2. Вычислим величину S_k, которая является ковариацией δz_k^- и суммой двух независимых случайных процессов, поскольку шум измерений V_k считается центрированным белым шумом, не коррелирующим с априорной оценкой погрешности вектора состояния системы. Как правило, параметр δz_k^- не является эргодическим, но в пределах ограниченного временного интервала и для определенного набора выборок он может быть аппроксимирован как эргодический сигнал. Из этого следует:

$$S_k = Cov\left(\delta z_k^-\right) = Cov\left(H_k \hat{x}_k + V_k - H_k \hat{x}_k^-\right) = Cov\left(H_k\left(\hat{x}_k - \hat{x}_k^-\right)\right) + cov\left(V_k\right) =$$

$$= H_k Cov\left(\hat{x}_k - \hat{x}_k^-\right) H_k^T + cov\left(V_k\right),$$

$$S_k = H_k P_k H_k^T + R_k.$$

Все параметры S_k с очевидностью определяются с использованием ФК. Кроме того, $Cov\left(\delta z_k^-\right)$ можно вычислить статистически с помощью вектора измерений – обозначим эту величину C_{Inn}. Тогда вознаграждение опишем так:

$$\text{Вознаграждение} = -\text{норма}(C_{Inn} - S_k). \tag{20}$$

Элемент C_{Inn} можно рассчитать следующим образом (для окна размерностью M величина $\tilde{z}$ является реальным измерением):

$$C_{Inn} = \frac{1}{M} \sum_{i=k-M+1}^{k} \left(\tilde{z}_i - H_i \hat{x}_i^-\right)\left(\tilde{z}_i - H_i \hat{x}_i^-\right)^T. \tag{21}$$

Эргодичность δz_k^- принимается в пределах установленного временного интервала, на котором БПЛА проявляет квазистационарную динамику. В течение этого периода изменения вектора состояния и шума измерений минимальны, что позволяет приблизительно оценить эргодичность. Например, во время стабильного полета с незначительными внешними возмущениями стохастические свойства δz_k^- имеют устойчивую динамику, благодаря чему этот параметр можно аппроксимировать как эргодический на длительных временных промежутках. При активном маневрировании нельзя производить расчеты на длительных временных интервалах из-за быстрых изменений динамики, приводящих к тому, что принятое допущение становится недействительным. Т.е. временные интервалы обработки измерений необходимо сокращать. Было установлено, что для периода выборки 0,01 с (100 выборок в секунду) наилучший размер окна M – 10 выборок. Величина M должна быть выбрана таким образом, чтобы обеспечить баланс между вычислительной эффективностью и эргодичностью. Так, при $M = 1$ возрастает сложность вычислений, а при слишком больших значениях M нарушается условие эргодичности. В результате для уравно-

вешивания условий достаточно, чтобы M = 5, 10 и 15. Стандартом здесь является векторная норма главной диагонали основания матрицы.

Вознаграждение (20) побуждает систему привести оценку ковариации в соответствие с фактической ковариацией, благодаря чему фильтр эффективнее справляется с неопределенностью R. Благодаря фокусировке на R вознаграждение (20) позволяет уменьшить влияние шума на оценку вектора состояния, тем самым дополняя влияние первого вознаграждения.

Очевидно, что возвращаемое значение (вознаграждение) противоположно погрешности, соответственно, максимизация вознаграждений означает минимизацию погрешностей. В других областях вознаграждение может вычисляться с помощью разных уравнений в зависимости от выбранной системы [20, 21].

Параметры Q-обучения

Стратегия Q-обучения в ОП характеризуется следующим [1]:

- состояние (S): в любой момент времени агент находится в определенном состоянии в среде. Состояние – это текущая ситуация или конфигурация среды; каждое состояние задается значением матрицы R с текущим Q-значением, и агент может перейти в это состояние из другого;
- действие (A): агент может совершать различные действия, находясь в определенном состоянии. Действия являются решениями о переходе агента из текущего состояния в новое, в результате чего формируется новое значение матрицы R. Для одних задач заранее могут быть установлены определенные действия, которые будут выполнять агенты, для других набор действий не ограничивается; в предлагаемом нами подходе – второй вариант;
- вознаграждение: после выполнения какого-либо действия и перехода в новое состояние (формирования нового значения матрицы R) агент получает от среды вознаграждение, которое показывает, насколько удачным было его действие в данном конкретном состоянии. Цель агента – получить как можно больше вознаграждений, при этом он может их накапливать или сбрасывать на каждой итерации;
- α – коэффициент обучения, устанавливающий, в какой степени новая информация определяется старой;
- η_d – коэффициент ценности, обозначающий уровень значимости будущих вознаграждений;
- QL_{sa} – ожидаемая величина накопленных вознаграждений, которые агент получает за конкретное действие в определенном состоянии при дальнейшем следовании оптимальной стратегии. По сути, это показатель качества действия, выполненного в заданном состоянии (высокое значение QL_{sa} означает большее вознаграждение за правильное действие);
- MAX_Q – максимальное значение QL_{sa} для следующего состояния S с учетом всех возможных действий A;
- NP – количество периодов или эпизодов.

Комбинации

Существуют две возможные комбинации, одна из которых рассматривается в [7, 8] и предполагает наличие обычного ОФК, обучающего или поискового ОФК и обу-

ченного ОФК, в котором матрица R является настраиваемым параметром. Эта комбинация представлена в табл. 4.

Таблица 4

Первая комбинация ОП и ОФК

Первая комбинация
Для периода в *NP* (эпизод): • для каждой выборки данных: ◁ применение нормального ОФК1 (R и другие параметры), ◁ применение поискового ОФК2 (R и другие параметры), ◁ расчет накопленных вознаграждений по результатам ОФК1 и ОФК2, ◁ определение наилучшего действия (формирование новой R^+), ◁ применение к обученному ОФК (R^+ и другие параметры); • завершение выборок. Завершение периодов.

Выходными данными каждого ОФК являются векторы состояния и измерений, а входными – векторы состояния и измерений, а также матрицы шума измерений и шума возмущений. При этом ОФК1 и ОФК2 – нормальные ОФК, но ОФК2 привлекается в текущем действии Q-обучения, а ОФК1 – для обработки результатов предыдущих действий. Это означает, что в выборке k:

- ОФК2 применяется для текущего действия;
- выходные данные (вектор состояния) ОФК2 в выборке k–1 служат в качестве входных (вектор состояния) для ОФК1 в выборке k;
- величина матрицы R вычисляется в соответствии с ОФК1 и ОФК2 и используется в составе входных данных для обученного ОФК;
- окончательные параметры ориентации получают из обученного ОФК.

Более подробную информацию можно найти в [8]. При традиционном ОП агент выполняет действия в необходимой последовательности в рамках одного эпизода и накапливает вознаграждения в зависимости от полезности каждого действия. Положительные вознаграждения увеличивают количество суммарных баллов, а отрицательные – уменьшают. Эпизод завершается, когда накопленные вознаграждения опускаются ниже пороговой величины, что свидетельствует о выборе R, адекватной внешним условиям, после этого начинается новый эпизод. Эпизоды с высокоценными вознаграждениями сохраняются с целью фиксации эффективных стратегий. В настоящей работе предлагается подход, не предусматривающий накопление вознаграждений – вместо этого оценивается вознаграждение, полученное за каждое действие на каждой итерации алгоритма ОФК. При этом в рамках одного эпизода проверяются все возможные действия, и за каждое из них вознаграждение рассчитывается отдельно. Действие, за которое было получено наибольшее вознаграждение, выбирается для следующего эпизода. На базе успешных эпизодов определяется последовательность оптимальных действий, которая применяется на следующей итерации ОФК.

Продемонстрируем отличительные особенности данного метода на примере игры в жанре «квест», в пространстве которой игрок может двигаться вправо, влево, вверх или вниз (ограниченные действия). Эпизод завершается, когда найдено «сокровище» или «взорвалась мина». В нашем случае эпизоды не имеют четкого завершения – они заканчиваются, когда выполняемое действие больше не влечет за собой изменение Q-значения нового состояния. Основная идея состоит в том, что действие либо при-

водит к новому состоянию, либо изменяет текущее состояние, либо сохраняет его. В конце эпизода действие, обусловившее состояние с максимальным Q-значением, выбирается для следующего эпизода. Эта комбинация (табл. 5) сложнее, но позволяет непрерывно обновлять ОФК за счет последовательного выбора действий, оптимизирующих Q-значение и повышающих эффективность оценивания углов ориентации. В рамках одного эпизода следующее действие выбирается из существующего набора случайным образом, поскольку при сохранении всех Q-значений таковое у текущего действия зависит от Q-значения предыдущего, вознаграждения и максимального Q-значения всех действий. Выбор осуществляется по формуле

$$QL_{s,a} = (1-\alpha)QL_{s,a} + \alpha\left(Reward + \eta_d \times MAX_Q\right). \tag{22}$$

Следующее действие между эпизодами выбирается исходя из максимальных Q-значений последнего эпизода.

Таблица 5

Предлагаемая новая комбинация *Q*-обучения и ОФК

Вторая комбинация
$Q_{\text{старое значение}} = 0$. $R^+ = 0$ – матрица того же размера, что и R. Для каждой выборки данных: • применение ОФК (матрица R и другие параметры); • для каждого периода из NP периодов (эпизод): ◁ совершение действия, ◁ расчет вознаграждения, ◁ вычисление Q-значения ($Q_{\text{старое значение}}$, текущее действие), ◁ $Q_{\text{старое значение}}$ = Q-значение наилучшего действия с самым большим вознаграждением; • выполнение действия с наилучшим вознаграждением (получение соответствующего значения матрицы R и его подстановка в матрицу R^+); • $Q_{\text{старое значение}}$ = Q-значение наилучшего действия в последнем эпизоде; • $R = R^+$; • завершение периодов. Завершение выборок.

Этапы процесса Q-обучения (выбор действия, расчет вознаграждения, вычисление Q-значения и другие шаги) приведены в табл. 6.

Таблица 6

Этапы процесса *Q*-обучения

Адаптация ОП
Для периода в NP (эпизод): • выбор индекса действия (ind) от 1 до NA (NA – количество действий); • определение уровня адаптации (A) в соответствии с индексом действия: $A = ind$ или, как вариант, $A = \beta \times ind$, где β – случайное действительное число в интервале [0–ind]; • адаптация матрицы R: $R = R + A \times RM$ (RM – случайная матрица одного размера с матрицей R); • расчет вознаграждения; • расчет максимального Q-значения для следующего состояния: $\text{Max}_Q = \text{Max}(QL_{s,a})$; • корректировка Q-значения с помощью специального уравнения: $QL_{s,a} = (1 - a)\, QL_{s,a} + a(Reward + \eta_d \times \text{Max}_Q)$; • выбор $QL_{s,a}$ для действия с наибольшим Q-значением, чтобы определить величину R.

Для каждого действия от 1 до *NA* генерируется новое значение *RM*, при этом уровень адаптации – $A \times RM$, где $A = \beta \times ind$, β – случайное действительное число в интервале [0–*ind*]. Указанная адаптивная часть сохраняется в памяти, а затем рассчитываются вознаграждение и *Q*-значение для данного действия в соответствии с (21). В конце эпизода каждое действие обладает *Q*-значением и адаптивной частью. Оценка действий по их *Q*-значениям позволяет выбрать такое из них, которое обеспечивает максимальное вознаграждение согласно принципам *Q*-обучения. Каждое *Q*-значение обновляется итеративно по формуле (21), учитывающей предыдущее *Q*-значение, полученное вознаграждение и максимальное *Q*-значение всех возможных действий. Описанная процедура итерационного обновления *R* дает возможность отдавать приоритет таким действиям, которые с течением времени повышают эффективность работы ОФК за счет адаптивной коррекции шума.

Предложенный метод предусматривает генерацию уникальной случайной матрицы *RM* для каждого действия в пределах одного эпизода, чтобы внести вариативность в процесс адаптации к шуму. Эта матрица служит основой для масштабирования матрицы ковариации шума *R* с помощью коэффициента адаптации и создает ряд возможных конфигураций шума для оценивания вектора состояния системы. Между эпизодами матрица *RM* генерируется снова, что позволяет алгоритму *Q*-обучения исследовать различные конфигурации шума в последовательных эпизодах. Такая непрерывная адаптация повышает способность алгоритма находить наиболее эффективные варианты корректировки модели шума с целью более точной оценки углов ориентации БПЛА.

4. Реализация

Для реализации нового метода были выбраны второй тип вознаграждения (уравнение (20)) и комбинация, описанная в табл. 5. Приведенные данные взяты из интернета. Они были получены в ходе двух разных экспериментов во время полетов БПЛА самолетного типа. Измерения выполнялись с помощью датчика MPU 9250 – девятиосного устройства, состоящего из трехосного гироскопа, трехосного акселерометра, трехосного магнитометра и процессора Digital Motion Processor™ (DMP), размещенных в небольшом корпус размером 3 × 3 × 1 мм. В качестве эталона использовалась информация об ориентации, полученная от интегрированной инерциально-спутниковой навигационной системы (ИНС-GPS) с частотой выборки 64 Гц. Начальное значение матрицы *R* задано в соответствии с руководством по эксплуатации датчиков. Входными данными для *Q*-обучения являются описанные выше параметры C_{Inn}, S_k, с помощью которых алгоритм *Q*-обучения находит наилучшие варианты адаптации матрицы *R*. Метод предназначен для применения в условиях неустойчивого сигнала GPS или его отсутствия.

Из рисунков следует, что метод ОП-ОФК превосходит традиционный ОФК. На укрупненных фрагментах диаграмм заметно, что в одних областях ОП-ОФК и обычный ОФК имеют схожую эффективность, а в других результаты ОП-ОФК лучше, чем у ОФК. В каждый конкретный момент времени формируется новая матрица *R*; это означает, что она адаптируется к различным типам движения – устойчивому полету и маневрированию. Результаты, полученные с помощью ОФК и ОП-ОФК, сопоставлены ниже в табл. 7.

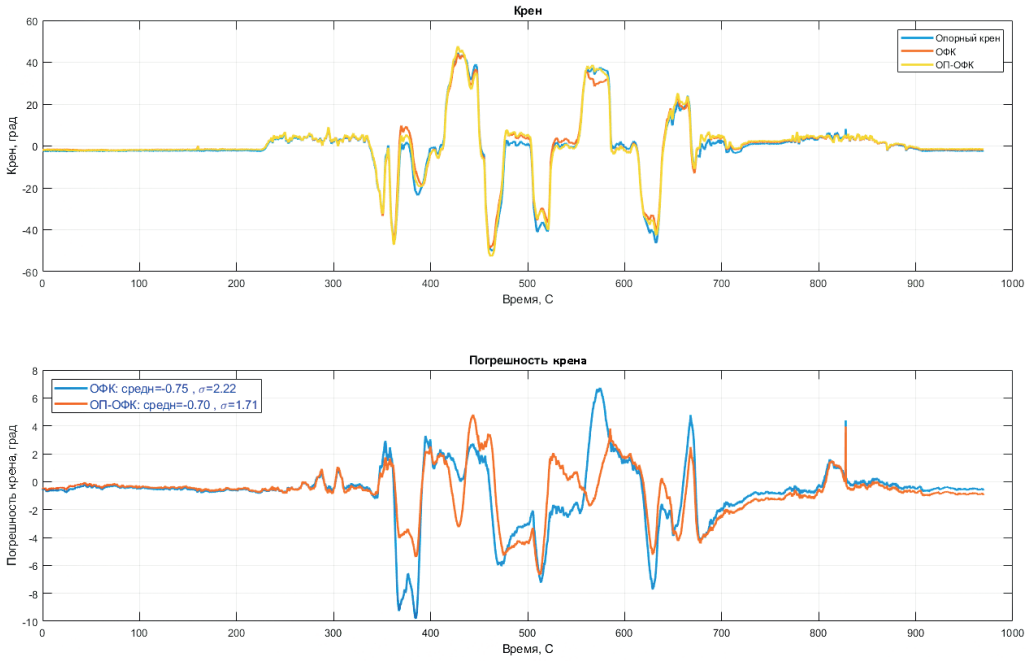

Рис. 1. Оценивание угла тангажа (эксперимент 1)

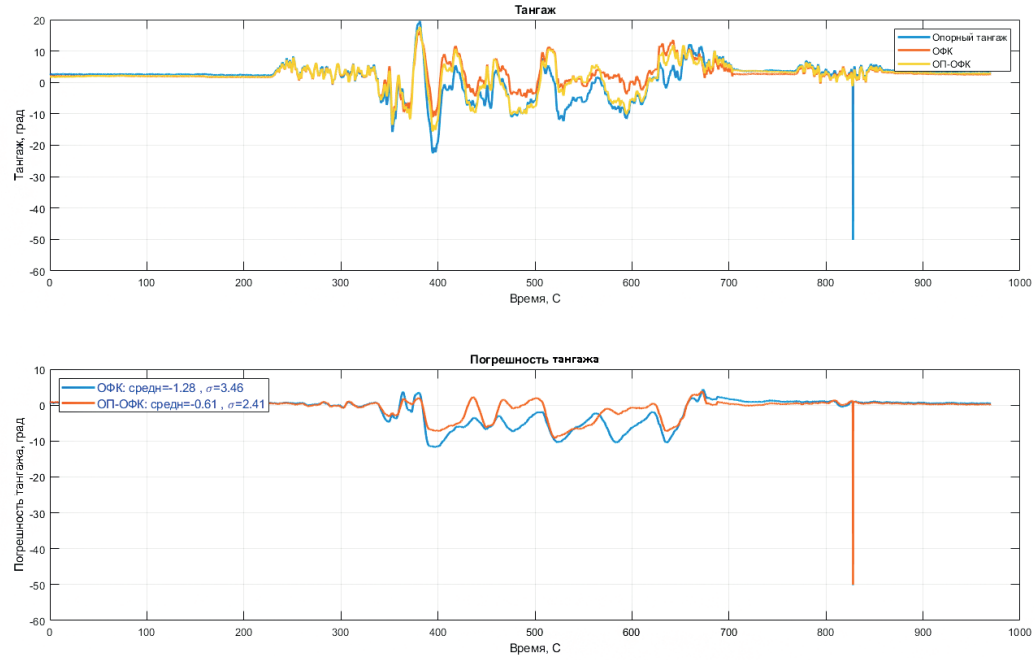

Рис. 2. Оценивание угла крена (эксперимент 1)

При сравнении было установлено, что метод на основе ОФК в сочетании с Q-обучением позволяет существенно повысить точность определения углов ориентации по сравнению с обычным ОФК. Новый подход может быть весьма полезным при неустановившихся полетах, когда матрица ковариации шума измерений не является статичной и требует адаптивной коррекции.

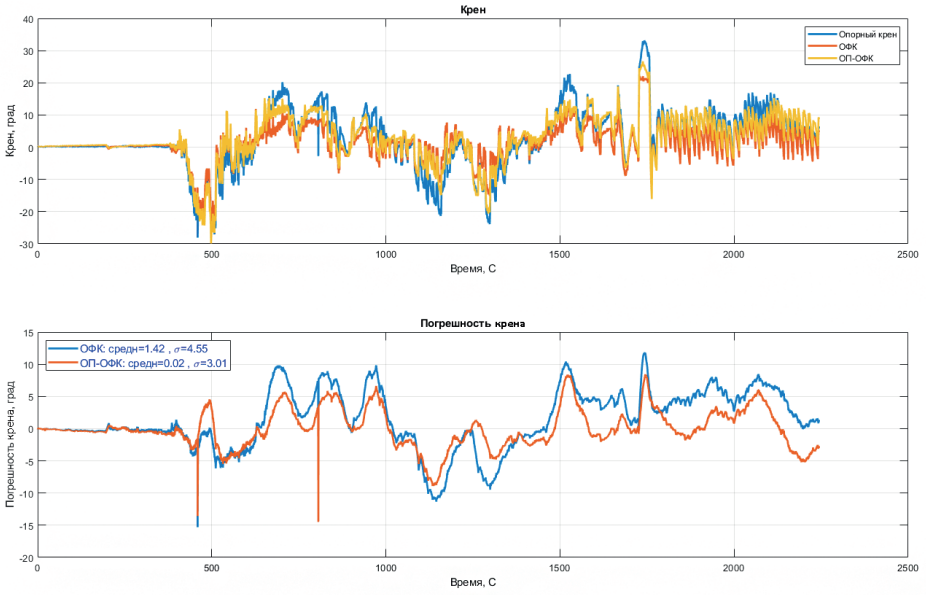

Рис. 3. Оценивание угла тангажа (эксперимент 2)

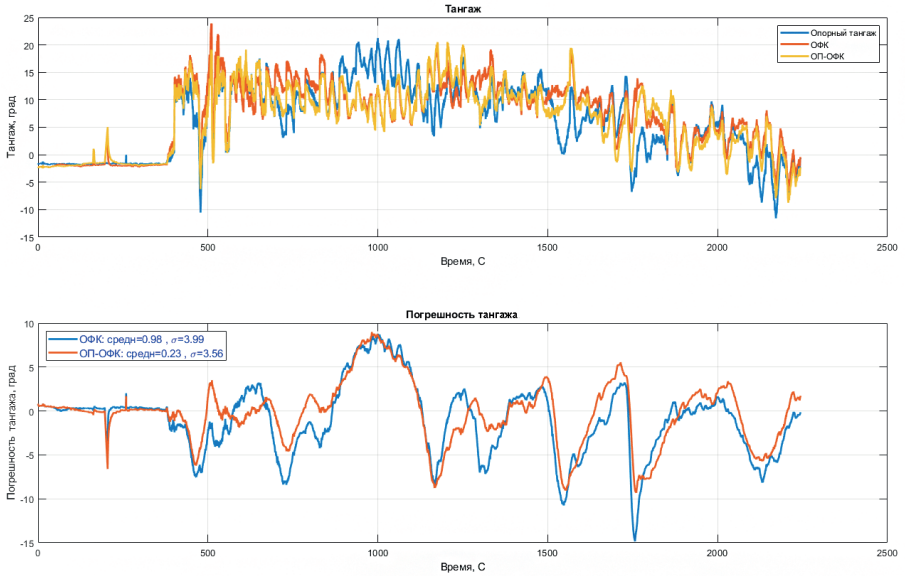

Рис. 4. Оценивание угла крена (эксперимент 2)

Таблица 7

Сравнение результатов

Погрешность	Среднеквадратическое отклонение (°)				Средняя (°)			
	Крен	Тангаж	Крен	Тангаж	Крен	Тангаж	Крен	Тангаж
	Эксп. 1	Эксп. 1	Эксп. 2	Эксп. 2	Эксп. 1	Эксп. 1	Эксп. 2	Эксп. 2
Только ОФК	3,46	2,22	3,98	4,55	1,28	0,75	0,98	1,42
ОП-ОФК (принцип *Q*-обучения)	2,41	1,71	3,56	3,01	0,61	0,70	0,23	0,02
Улучшение	30,35%	22,97%	10,55%	33,85%	52,34%	5,33%	76,53%	98,59%

Заключение

В настоящей работе предложен метод с использованием обучения с подкреплением, основанного на принципе *Q*-обучения, позволяющий повысить точность ОФК при оценивании углов ориентации БПЛА. Показано, что ОФК в сочетании с ОП более эффективен, чем традиционный ОФК, что подтверждает потенциальную возможность реализации нового метода в системах навигации БПЛА.

ЛИТЕРАТУРА

1. **Naeem, M., Rizvi, S. T. H., and Coronato, A.**, A gentle introduction to reinforcement learning and its application in different fields, *IEEE Access*, 2020, vol. 8, pp. 209320–209344.
2. **Du, W., and Ding, S.**, A survey on multi-agent deep reinforcement learning: from the perspective of challenges and applications, *Artificial Intelligence Review*, 2021, vol. 54, no. 5, pp. 3215–3238.
3. **Goslinski, J., Giernacki, W., and Krolikowski, A.**, A nonlinear filter for efficient attitude estimation of unmanned aerial vehicle (UAV), *Journal of Intelligent & Robotic Systems*, 2019, vol. 95, pp. 1079–1095.
4. **Jing, X., Cui, J., He, H., Zhang, B., Ding, D., and Yang, Y.,** Attitude estimation for UAV using extended Kalman filter, *Proc. 2017 29th Chinese Control And Decision Conference (CCDC)*, May 2017, pp. 3307–3312.
5. **Assad, A., Khalaf, W., and Chouaib, I.,** Novel adaptive fuzzy extended Kalman filter for attitude estimation in GPS-denied environment, *Gyroscopy and Navigation*, 2019, vol. 10, no. 3, pp. 131–146.
6. **Crassidis, J. L.,** Spacecraft attitude determination, in *Encyclopedia of Systems and Control,* Cham: Springer International Publishing, 2021, pp. 2097–2104.
7. **Xiong, K., Wei, C., and Zhang, H.,** Q-learning for noise covariance adaptation in extended Kalman filter, *Asian Journal of Control*, 2021, vol. 23, no. 4, pp. 1803–1816.
8. **Dai, X., Nateghi, V., Fourati, H., and Prieur, C.,** Q-learning-based noise covariance adaptation in Kalman filter for MARG sensors attitude estimation, *Proc. 2022 IEEE International Symposium on Inertial Sensors and Systems (INERTIAL),* May 2022, pp. 1–6.
9. **Pandey Y., Bhattacharyya R., Nath Singh Y.** Robust Attitude Estimation with Quaternion Left-Invariant EKF and Noise Covariance Tuning, *arXiv e-prints*, 2024, pp. arXiv: 2409.11496.
10. **Odry, A., Fuller, R., Rudas, I. J., and Odry, P.,** Kalman filter for mobile-robot attitude estimation: Novel optimized and adaptive solutions, *Mechanical Systems and Signal Processing*, 2018, vol. 110, pp. 569–589.
11. **Odry, A., Kecskes, I., Sarcevic, P., Vizvari, Z., Toth, A., and Odry, P.,** A novel fuzzy-adaptive extended Kalman filter for real-time attitude estimation of mobile robots, *Sensors*, 2020, vol. 20, no. 3, p. 803.
12. **Hajiyev, C., and Soken, H. E.,** Robust adaptive unscented Kalman filter for attitude estimation of pico satellites, *International Journal of Adaptive Control and Signal Processing*, 2014, vol. 28, no. 2, pp. 107–120.
13. **Qiu, Y., Su, Y., and He, X.,** An improved IMM-KF for UAV position prediction, *Proc. Third International Conference on Control and Intelligent Robotics (ICCIR 2023)*, December 2023, vol. 12940, pp. 7–16.
14. **Parwana, H., and Kothari, M.,** Quaternions and attitude representation, *arXiv preprint arXiv:1708.08680,* 2017.
15. **Ribeiro, M. I.,** Kalman and extended Kalman filters: Concept, derivation and properties, *Institute for Systems and Robotics*, 2004, vol. 43(46), pp. 3736–3741.
16. **Quan, W., Li, J., Gong, X., and Fang, J.,** *INS/CNS/GNSS Integrated Navigation Technology,* Springer, 2015.
17. **Chouaib, A.I., Wainakh, B.M., and Khalaf, C.W.**, Robust self-corrective initial alignment algorithm for strap-down INS, *Proc. 10th Asian Control Conference (ASCC)*, 2015, pp. 1–6.
18. **Trawny, N., Roumeliotis, S.I.,** *Indirect Kalman filter for 3D attitude estimation,* University of Minnesota, Dept. of Comp. Sci. & Eng., Tech. Rep., 2005, vol 2, p. 2005.
19. **Center, N.G.D.,** NCEI geomagnetic calculators [Электронный ресурс]. URL: https://www.ngdc.noaa.gov/geomag/calculators/magcalc.shtml (дата обращения: 01.04.2025).
20. **Wang, J., Liu, Y., and Li, B.,** Reinforcement learning with perturbed rewards, *Proc. of the AAAI Conference on Artificial Intelligence,* 2020, vol. 34, no. 4, pp. 6202–6209.
21. **Icarte, R. T., Klassen, T. Q., Valenzano, R., and Mcllraith, S. A.,** Reward machines: Exploiting reward function structure in reinforcement learning, *Journal of Artificial Intelligence Research*, 2022, vol. 73, pp. 173–208.

Assad, A., Serikov, S.A. (St. Petersburg State University of Aerospace Instrumentation)
Adaptation Noise Covariance in Extended Kalman Filter Using Reinforcement Learning for Improved UAV Attitude Estimation, *Giroskopiya i Navigatsiya*, 2025, vol. 33, no. 3 (130), pp. 33–50.

Abstract. Accurate attitude determination of Unmanned Aerial Vehicles (UAVs) is crucial for autonomous navigation, particularly when relying solely on gyroscope, accelerometer, and magnetometer measurements without utilizing the Global Positioning System (GPS). Reinforcement Learning (RL) has emerged as a promising artificial intelligence technique applicable across various domains. This research introduces a novel approach that leverages RL to enhance the performance of the Extended Kalman Filter (EKF) in attitude estimation. The proposed method depends of RL which uses the Q-Learning model and policy to find best solution to adjust autonomously the measurement noise covariance matrix within the EKF. By establishing a reward mechanism that incentivizes actions minimizing the prediction error relative to true measurements, RL dynamically optimizes the measurement noise covariance matrix. This innovative integration of RL and EKF, referred to as RL-EKF, has been implemented and tested. Results demonstrate that RL-EKF significantly outperforms the traditional EKF, yielding marked improvements in attitude estimation accuracy, the improvement ratios showed that selected method is very effective in the field of attitude estimation.

Key words: Unmanned Aerial Vehicle, Reinforcement Learning, Extended Kalman filter, Attitude estimation.

Материал поступил 20.08.2024