

Разработка гибридной нейронной сети глубокого обучения с использованием квадратнокорневого сигма-точечного фильтра Калмана для оценки массы автомобиля и уклона дороги

Хоу Тяньюй, А.А. Кулик

Московский Государственный Технический Университет им Н.Э. Баумана, Москва

Аннотация: В статье представлена гибридная нейронная сеть оценки массы автомобиля и продольного/поперечного уклонов дороги, объединяющая квадратнокорневой сигма-точечный фильтр Калмана и нейросетевую модель на базе энкодера трансформера с использованием перекрёстного внимания к остаткам оценивания. Предложенный подход сочетает физическую интерпретируемость фильтра с высокой аппроксимационной способностью нейросети. Для обеспечения внедрения на встраиваемых электронных блоках управления модель была упрощена с помощью преобразования знаний в компактную сеть долгой краткосрочной памяти. Результаты экспериментов в различных сценариях показали снижение средней ошибки более чем на 25 % при вычислительной задержке менее 0,3 мс.

Ключевые слова: оценка состояния транспортного средства, оценка уклона дороги, оценка массы автомобиля, нейросеть типа трансформер, перекрёстное внимание, адаптивная фильтрация, дистилляция знаний, квадратнокорневой сигма-точечный фильтр Калмана, интеллектуальные транспортные средства, слияние датчиков.

Введение

Современные методы оценки динамических параметров транспортного средства условно подразделяются на две категории: модельно-ориентированные и основанные на данных. Модельно-ориентированные алгоритмы используют динамические или кинематические модели автомобиля совместно с фильтрами типа Калмана или нелинейными наблюдателями. Недавние работы показывают: адаптивные фильтры Калмана с двойными факторами забывания для онлайн-настройки ковариаций шумов процесса и измерений (Q и R) [1,2]; двухуровневую фильтрацию для условий отсутствия глобальной навигационной спутниковой системы при поддержке инерциального измерительного модуля [3,4]; режимное переключение для сцен пробуксовки колёс [5]. Наблюдатель-ориентированные подходы, включая алгоритмы скользящего режима и гибриды рекурсивного метода наименьших квадратов с расширенным фильтром Калмана [6–9] снижают

требования к датчикам, но чувствительны к рассогласованию модели и пороговой логике. Многомодельные методы с объединением доказательств повышают надёжность, однако требуют существенно больших вычислительных затрат [10].

Подходы, основанные на данных, используют методы глубокого обучения для моделирования многомерных временных рядов. Свёрточные и рекуррентные нейронные сети, а также облегчённые архитектуры на основе управляемых рекуррентных блоков, повышают точность краткосрочного прогнозирования [11,12]. Вместе с тем такие модели нередко оказываются недостаточно физически согласованными и могут хуже обобщаться при изменении условий движения.

Оценивание поперечного уклона дороги остаётся сложной задачей. Методы компьютерного зрения деградируют при неблагоприятной видимости (заслонения, тени, блики), а применение лидара требует дорогостоящих датчиков [13]. Для сочетания интерпретируемости оценок, получаемых методами фильтрации, и адаптивности нейросетевых моделей предложены гибридные схемы, в которых нейронная сеть динамически корректирует параметры фильтра Калмана [14].

Модель динамики транспортного средства и расчёт сил шин

Модель динамики транспортного средства с семью степенями свободы включает три степени свободы корпуса – продольную скорость v_x , поперечную скорость v_y и угловую скорость рыскания ω , и четыре угловые скорости вращения колёс (рис.1 и рис.2). На основе модели учитывается динамика подвески (тангаж/крен относительно дороги) для восстановления истинных углов корпуса и уклонов дорожного полотна.

Уравнения движения имеют следующий вид [15]:

$$\begin{aligned}\dot{\omega} &= \frac{1}{I_z} \left\{ [(F_{Xf1} + F_{Xf2}) * \sin(\delta_f) + (F_{Yf1} + F_{Yf2}) * \cos(\delta_f)] * a + \frac{l_1}{2} \right. \\ &\quad * [(F_{Xf1} - F_{Xf2}) * \cos(\delta_f) + (F_{Yf1} - F_{Yf2}) * \sin(\delta_f)] - \frac{l_2}{2} * (F_{Xr1} - F_{Xr2}) \\ &\quad \left. - (F_{Yr1} + F_{Yr2}) * b \right\} \\ a_x &= \frac{[-(F_{Xf1} + F_{Xf2}) * \sin(\delta_f) - (F_{Yf1} + F_{Yf2}) * \cos(\delta_f)]}{m} \\ a_y &= [(F_{Xf1} + F_{Xf2}) * \cos(\delta_f) - (F_{Yf1} + F_{Yf2}) * \sin(\delta_f)]/m\end{aligned}\quad (1)$$

где m – масса автомобиля, I_z – момент инерции относительно вертикальной оси, l_1, l_2 – ширины колеи переднего и заднего мостов, a, b – расстояния от центра масс до переднего и заднего мостов, δ_f – угол поворота передних колёс.

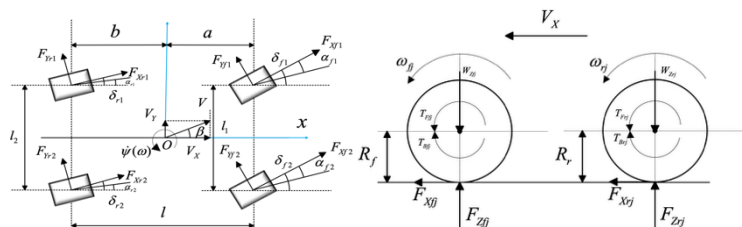


Рис. 1 – Модель динамики транспортного средства с 7-степенями свободы

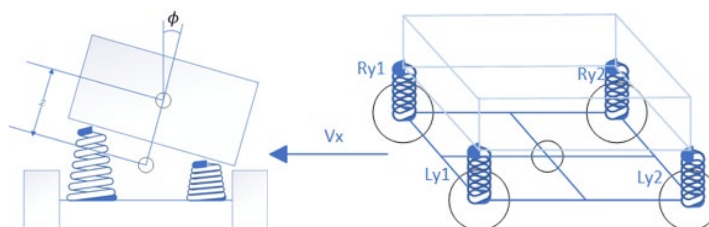


Рис. 2. – Модель крена с учётом степени свободы подвески.

Под влиянием продольных/поперечных ускорений происходит перераспределение вертикальных нагрузок; для переднего левого колеса:

$$F_{Zf1} = \left(\frac{1}{2} mg - ma_y * \frac{h_g}{l_1} \right) * \frac{b}{l} - \frac{1}{2} ma_x \left(\frac{h_g}{l} \right) + mgsin\gamma_x * \left(\frac{b}{l} \right) + mgsin\gamma_y \left(\frac{l_1}{h_g} \right), \quad (2)$$

Аналогичные выражения применяются для F_{Zf2} , F_{Zr1} и F_{Zr2} . Здесь h_g – высота центра масс, γ_x, γ_y – продольный и поперечный уклоны дороги. Для разделения влияния уклонов и положения кузова в вектор наблюдений добавляют датчики прогиба подвески (активная/полуактивная подвеска). При малых углах:

$$\begin{aligned}h_{front} &= (Ly1 + Ry1)/2 & h_{rear} &= (Ly2 + Ry2)/2 \\h_{left} &= (Ly1 + Ly2)/2 & h_{right} &= (Ry1 + Ry2)/2 \\Pitch_{body} &= \text{atan}\left(\frac{h_{front}-h_{rear}}{L}\right) & Roll_{body} &= \text{atan}\left(\frac{h_{right}-h_{left}}{l_1}\right)\end{aligned}\quad (3)$$

Такое геометрическое разделение позволяет более точно оценить уклоны дороги, отделяя их от движений кузова. Используя данные инерциального измерительного модуля и рассчитанные вертикальные нагрузки, можно оценить боковые силы шин согласно аппроксимирующей модели (формула Пачейки) [16]:

$$F_Y(\alpha) = D * \sin\{C * \arctan[B * \alpha - E(B * \alpha - \text{atan}(B * \alpha))]\} + Sv, \quad (4)$$

где α – угол скольжения, а B, C, D, E – коэффициенты формы. Они подогнаны по данным CarSim и уточнены дорожными испытаниями. С учётом зависимости от вертикальной нагрузки:

$$\begin{aligned}F_Y(\alpha) &= (a_1 F_Z^2 + a_2 F_Z + a_3) * \\&* \sin\{C * \arctan[B * \alpha - E(B * \alpha - \text{atan}(B * \alpha))]\} + (Sv * F_Z + a_4),\end{aligned}\quad (5)$$

Адаптивный квадратнокорневой сигма-точечный фильтр Калмана для совместной оценки массы автомобиля и уклонов дороги

В процессе выполнения работы рассматривается оценка трёх основных параметров: массы автомобиля и продольного (γ_x) и поперечного (γ_y) уклонов дороги. Вектор состояния имеет следующий вид:

$$x = [lnm, \gamma_x, \gamma_y]^T, \quad (6)$$

где lnm – элемент стабилизации оценивания, подавления мультипликативного шума и повышения устойчивости.

Учитывая, что изменения уклонов дороги малы, их динамика моделируется с использованием метода «блуждания»:

$$x_k = x_{k-1} + w_{k-1} \quad w_{k-1} \sim N(0, Q), \quad (7)$$

Вектор наблюдения включает продольное и поперечное ускорения a_x , a_y и оценки уклонов получены из ИНС и датчиков положения кузова

транспортного средства. Вектор наблюдения может быть представлен выражением:

$$\begin{aligned} Z_k &= [a_x, a_y, \gamma_x, \gamma_y]^T h(x_k, u_k) + v_k, \\ a_x &= \frac{F_x}{m} + g \sin \gamma_x, \quad a_y = \frac{F_y}{m} - g \sin \gamma_y, \\ \gamma_x &= \gamma_{xIMU} - \gamma_{xbody}, \quad \gamma_y = \gamma_{yINU} - \gamma_{ybody}, \quad m = \exp(x_1) \end{aligned} \quad (8)$$

где γ_{xIMU} - угол продольного наклона, измеренный ИНС; γ_{xbody} - угол продольного наклона кузова транспортного средства относительно поверхности дороги; γ_{yINU} - угол поперечного наклона, измеренный ИНС, γ_{ybody} - угол поперечного наклона кузова транспортного средства относительно поверхности дороги.

Из-за шумов датчиков (например, подвески) требуется фильтрация наблюдений. В алгоритме квадратнокорневого сигма-точечного фильтра Калмана (SR-UKF) масса инициализируется номинально, углы уклона — нулём; ковариации процесса Q и измерений R разлагают по Холецкому. Для размерности состояния $n = 3$ формируются $2n + 1$ сигма-точек:

$$\begin{cases} x_k^0 = \hat{x}_k \\ x_k^i = \hat{x}_k + \sqrt{n + \lambda} * S_k \quad i = 1 \dots n, \\ x_k^{i+n} = \hat{x}_k - \sqrt{n + \lambda} * S_k \quad 1 \dots n \end{cases} \quad (10)$$

Здесь λ — параметр масштабирования, определяющий разброс сигма-точек вокруг среднего значения, n — размерность вектора состояния. Выражение $\sqrt{n + \lambda} * S_k$ представляет собой отклонение от среднего. Параметр λ задаётся в виде уравнения:

$$\lambda = \alpha^2(n + \kappa) - n, \quad (11)$$

где α определяет разброс сигма-точек (меньшие значения — более плотное распределение), а κ служит для улучшенного учёта высших моментов; веса сигма-точек используются при вычислении взвешенного среднего и ковариаций.

Веса для каждой сигма-точки определяются согласно выражению:

$$W_0^m = \frac{\lambda}{n + \lambda} \quad W_0^c = \frac{\lambda}{n + \lambda} + (1 - \alpha^2 + \beta)$$

$$W_i^m = W_i^c = \frac{1}{2^{*(n+\lambda)}}, i = 1, \dots, 2n, \quad (12)$$

Каждая сигма-точка проходит через нелинейную модель процесса, и предсказанное среднее вычисляется по формуле:

$$\hat{X}_{k+1|k} = \sum W_i^m X_{k+1|k}^{(i)}, \quad (13)$$

Остатки вычислений имеют вид:

$$\chi_i = X_{k+1|k}^{(i)} - \hat{X}_{k+1|k}, \quad (14)$$

Корни ковариации состояния получены на основе объединения остатков χ_i с $\sqrt{Q_k}$ с дальнейшим разложением с помощью QR-факторизации:

$$S_{x,k+1|k} = QR([\sqrt{W_i^c} * \chi_i, \sqrt{Q_k}]), \quad (15)$$

Прогнозирование измерений при размерности, равной четырем имеет вид:

$$Z_{k+1|k}^{(i)} = h(X_{k+1|k}^{(i)}) \quad \hat{Z}_{k+1|k} = \sum_{i=1}^{2n} W_i^m * Z_{k+1|k}^{(i)}, \quad (16)$$

Остатки измерений и их ковариация представлены выражениями:

$$\xi_i = Z_{k+1|k}^{(i)} - \hat{Z}_{k+1|k}, \quad S_{z,k+1|k} = QR([\sqrt{W_i^c} * \xi_i, \sqrt{R_k}]), \quad (17)$$

Кросс-ковариация, усиление Калмана и обновление состояния имеет вид:

$$P_{xz,k+1|k} = \sum_{i=1}^{2n} W_i^c \chi_i \xi_i^T, k = P_{xz,k+1|k} (S_{z,k+1|k} S_{z,k+1|k}^T)^{-1} = P_{xz,k+1|k} S_{z,k+1|k}^{-T} S_{z,k+1|k}^{-1}$$

$$X_{k+1} = \hat{X}_{k+1|k} + Kk * \widetilde{y}_k, \quad (18)$$

Прогнозирование с фактической ковариацией определяются по формуле:

$$\widetilde{y}_k = z_k - \hat{Z}_{k+1|k}$$

$$S_k^{pred} = S_{z,k+1|k} S_{z,k+1|k}^T, \hat{S}_k = \widetilde{y}_k \widetilde{y}_k^T, \quad (19)$$

На основе выражения (19) получено адаптивное обновление матрицы R :

$$R_{k+1} = R_k(1 - \eta_R) + \eta_R [\hat{S}_k - S_k^{pred}]_{pos}, \quad (21)$$

где $\eta_R \in (0,1]$ – коэффициент адаптации шума измерений, $\eta_R [\hat{S}_k - S_k^{pred}]_{pos}$ – симметризует аргумент и обеспечивает его положительную полуопределённость. Если $\hat{S}_k - S_k^{pred} > 0$, фильтр недооценил шум измерений, и R увеличивается; иначе – уменьшается.

Адаптивное обновление матрицы Q имеет вид:

$$\Delta Q_k = Kk[\hat{s}_k - s_k^{pred}]Kk^T, Q_{k+1} = Q_k(1 - \eta_Q) + \eta_Q \Delta Q_k, \quad (22)$$

При $\eta_Q \in [0.001, 0.01]$, выбранных малыми для предотвращения расходимости. На Q_{k+1} накладываются верхние и нижние границы для поддержания численной устойчивости.

Обновление корня ковариации определяется по формуле:

$$U_k = KkS_{z,k+1|k}, S_{k+1} = chol(S_{x,k+1|k}S_{x,k+1|k}^T - U_k U_k^T), \quad (23)$$

где *chol* - разложение Холецкого для положительно определённой матрицы. Выражение (23) завершают адаптивный цикл фильтрации в SR-UKF, обеспечивая согласованное обновление Q и R при сохранении численной устойчивости корня ковариации.

Оценивание параметров движения автомобиля с использованием глубокой нейронной сети на основе трансформера с физическими ограничениями и многоголовым механизмом внимания

В работе для оценки параметров движения автомобиля использовалась нейросетевая модель типа «трансформер» с энкодерной структурой, оптимизированная для регрессии состояний по данным датчиков в скользящем временном окне. Архитектура включает позиционное кодирование, механизм самовнимания с четырьмя параллельными подпространствами обработки и три последовательно соединённых энкодерных блока. В каждом блоке применяется нормализация, остаточные связи и регуляризация, а также полносвязная подсеть размерности 256–1024–256. Выход скрытого представления преобразуется полносвязным слоем в десять оцениваемых параметров: массу автомобиля, продольный и поперечный уклоны дороги, продольную и поперечную скорости, углы тангажа, крена и рыскания, а также угол бокового увода. Для оценки вычислительной эффективности модель тестировалась на целевой вычислительной платформе с использованием ускоренных режимов

вычислений на графическом процессоре; полученная задержка на одну итерацию не превышает 0,3 мс.

Нормализованная последовательность $X \in R^{T \times D}$ с позиционными кодировками поступает в признаковое пространство трансформера. Многоголовое самовнимание формирует несколько «голов», работающих в разных линейных подпространствах для определения разнообразных нелинейных связей. Структура и процесс алгоритма трансформера представлены на рис.3 [17].

Для реализации внимания входная последовательность линейно отображается в три набора признаков — запросы (Q), ключи (K) и значения (V), после чего вычисляются коэффициенты внимания и формируется уточнённое представление последовательности. Эти операции описываются выражениями:

$$Q_i = XW_i^Q; K_i = XW_i^K; V_i = XW_i^V, \quad (24)$$

где X - входная последовательность размерности, и W_i^Q, W_i^K, W_i^V - матрицы весов для формирования векторов запросов Q_i , ключей K_i и значений V_i [17]. Выбор данной архитектуры обусловлен её способностью выделять информативные фрагменты временного окна и устойчиво работать при разнородных датчиках.



Рис. 3. – Структура и процесс алгоритма трансформер

Эти проекции формируют матрицы запросов (Q), ключей (K) и значений (V). Матрица Q кодирует текущие сенсорные данные (ускорения, углы, ходы

подвески, скорости и т.д.), отражающие динамическое состояние автомобиля. Матрица K задаёт контекст временного окна и обеспечивает поиск релевантных зависимостей, тогда как матрица V хранит преобразованные признаки для восстановления целевых параметров. Механизм внимания вычисляет корреляции между Q и K , формируя уточнённые представления на основе V , которые далее используются для регрессии массы автомобиля, уклонов дороги и динамических состояний. Оценки внимания между оцениваемыми параметрами вычисляются по формуле [17]:

$$A_i = \frac{Q_i K_i^T}{\sqrt{d_h}}, \quad (25)$$

где d_h - размерность соответствующего подпространства.

На вход алгоритма подаются ускорения по трем осям, углы крена, тангажа и рыскания; четыре хода подвески; скорости колёс; тяговые/тормозные моменты; углы руления. Физические ограничения (например, декуплирование ориентации кузова и уклона по данным подвески) обеспечивают согласованность оценок при высокой динамике, такой как резкие изменения продольного и поперечного ускорений, дорожного уклона и угла крена автомобиля.

Построение гибридной схемы оценки параметров движения автомобиля на основе энкодера трансформера и SR-UKF

Преимуществом применения SR-UKF алгоритма является улучшение точности и стабильности оценивания, медленно меняющихся состояний объекта моделирования (масса, уклон). Однако представление неровностей поверхности движения в виде шума процесса движения автомобиля приводит к временным задержкам при частом проявлении неровностей. Нейросетевая модель на базе энкодера трансформера лучше аппроксимирует нелинейности и обеспечивает более высокое быстродействие по сравнению с фильтром SR-

UKF. Однако в квазистационарных режимах такая модель может усиливать измерительный шум, что ухудшает качество оценивания.

Для объединения достоинств обоих подходов предложена гибридная схема оценивания параметров движения транспортного средства на основе энкодера трансформера и SR-UKF, имеющая следующий вид:

$$x_k^{Final} = \alpha_k x_k^{TR} + (1 - \alpha_k) x_k^{UKF}, \quad (26)$$

где x_k^{TR} - вектор состояния, оцененный нейросетевой моделью на базе энкодера трансформера, x_k^{UKF} - вектор состояния, вычисленный с помощью SR-UKF.

Механизм перекрёстного внимания использует текущий 30-мерный вектор измерений в качестве входа, а также информацию об остатках (разностях) между оценками, полученными SR-UKF и нейросетевой моделью, и эталонными значениями. Это позволяет по данным и по динамике остатков адаптивно перераспределять доверие к каждому оценщику параметров транспортного средства. Явное использование остатков обеспечивает гибкий баланс между устойчивостью фильтра SR-UKF и быстродействием нейросетевой модели, повышая точность оценивания в переходных режимах и снижая влияние шума в квазистационарных режимах. Кроме того, такая схема делает процесс объединения оценок более интерпретируемым, поскольку вклад каждого оценщика на каждом шаге вычислений можно отслеживать по весам внимания.

Выход перекрёстного внимания задаётся выражением:

$$Attn(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_{model}}}\right), \quad (27)$$

Обучение кросс-модельной сети для оценки параметров движения транспортного средства содержит следующие этапы:

1. Крупная нейросетевая модель обучается на входных временных окнах длиной 512 шагов. На вход подаются остатки (разности) между оценками,

полученными SR-UKF и нейросетевой моделью, и эталонными значениями. Функция потерь задаётся выражением:

$$\mathcal{L}_{Teacher} = \frac{1}{N} \sum_{k=1}^N \|\alpha_k x_k^{TR} + (1 - \alpha_k) x_k^{UKF} - x_k\|^2, \quad (28)$$

Задачей выполнения первого этапа обучения является получение оптимального веса сети α_k , который обеспечивает согласование выходов переменных алгоритмов SR-UKF и нейросетевой моделью при различных дорожных условиях движения транспортного средства. При этом размер модели увеличивается для повышения способности к изучению сложных корреляций между остатками сигма-точечного фильтра Калмана и уверенностью оценивания параметров.

2. Для ресурсно-ограниченных платформ выполняется перенос знаний от модели-преподавателя в компактную трёхслойную рекуррентную сеть типа долговременной краткосрочной памяти (по 512 скрытых единиц в слое). Поскольку модель-преподаватель использует остатки SR-UKF, которые недоступны при онлайн-работе, компактная сеть обучается воспроизводить выходы «преподавателя» без явного доступа к этим остаткам [18]. На вход подаётся 36-мерный вектор измерений, на выходе формируются веса α_k . Функция потерь имеет вид:

$$\mathcal{L}_{Student} = \frac{1}{N} \sum_{k=1}^N \|\alpha_k^{Student} x_k^{TransEnc} + (1 - \alpha_k^{Student}) x_k^{UKF} - \hat{x}_k^{Teacher}\|^2, \quad (29)$$

Здесь $\hat{x}_k^{Teacher}$ — слитый выход обученной модели-учителя. Такой подход сохраняет точность нейросетевой системы при существенном снижении вычислительной сложности вывода; ускорение обучения достигается за счёт использования смешанной точности вычислений.

Моделирование оценки параметров движения транспортного средства

Данные динамики автомобиля сгенерированы на платформе CarSim + MATLAB на протяжённых подъёмах/спусках, горных дорогах и городских маршрутах с частыми поворотами, что дало 500 000 образцов для обучения.

Рис. 4 показывает траекторию и продольные/поперечные ускорения, подтверждая высокую динамичность условий.

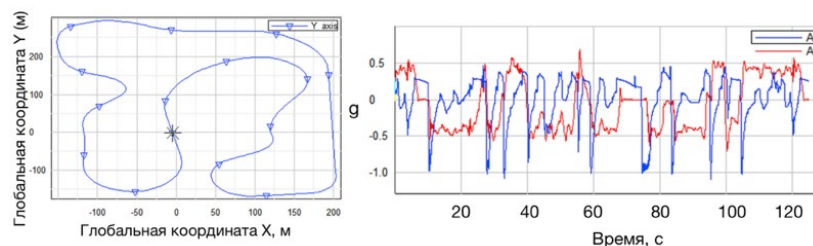


Рис. 4. Траектория имитационного испытания; продольные и поперечные ускорения на всём тестовом наборе.

Результаты моделирования при работе SR-UKF в офлайн-режиме представлены на рис. 5–7. Результаты моделирования показывают, что в установившихся (стационарных) режимах движения автомобиля ошибка оценки массы не превышает ± 50 кг. В переходных режимах ошибка может достигать 200 кг при интенсивном разгоне и пробуксовке, большие перемещения подвески ухудшают оценку уклонов. Оценка поперечного уклона запаздывает и ослабляется из-за нелинейных эффектов крена и скольжения.

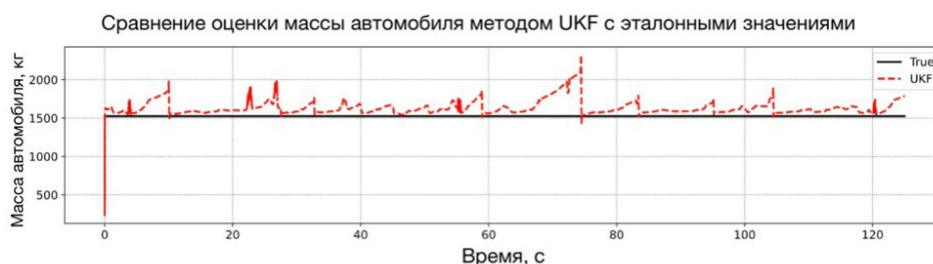


Рис. 5. – Сравнение оценки массы автомобиля с эталонными значениями.

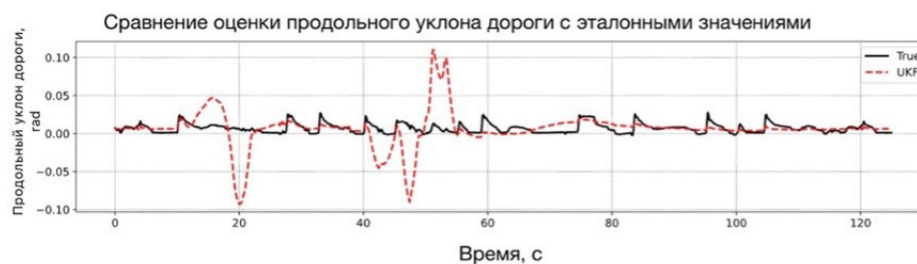


Рис. 6. – Сравнение оценки продольного уклона дороги с эталонными значениями

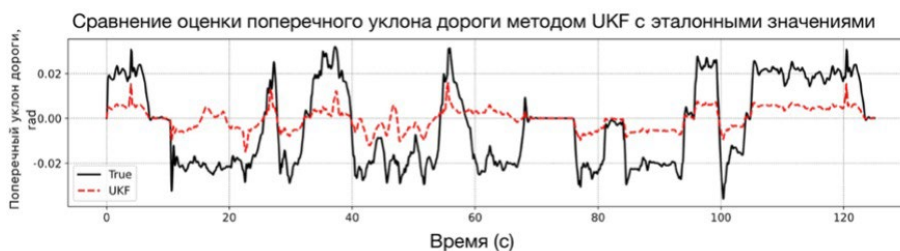


Рис. 7. – Сравнение оценки поперечного уклона дороги с эталонными значениями

Далее используются обозначения, принятые в тексте: вариант А - модель на базе энкодера трансформера (на рисунках: TransEnc), вариант В - рекуррентная модель с механизмом внимания (на рисунках: LSTM-Attn), вариант С - модель с механизмом перекрёстного внимания (на рисунках: CrossAttn). Моделирование показало, что среднеквадратичная ошибка оценки продольного уклона не превышает 0,02 рад, а поперечного уклона находится в диапазоне 0,01–0,02 рад. Среди рассмотренных нейросетевых вариантов наилучший результат обеспечивает модель с введёнными физическими ограничениями: среднеквадратичная ошибка не превышает 0,015 рад.

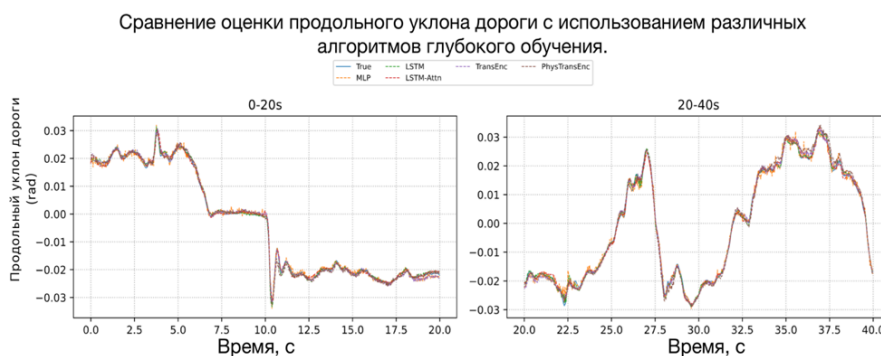


Рис. 8. – Сравнение оценки продольного уклона дороги с использованием различных алгоритмов глубокого обучения.

Сравнение оценки поперечного уклона дороги на основе различных моделей глубокого обучения

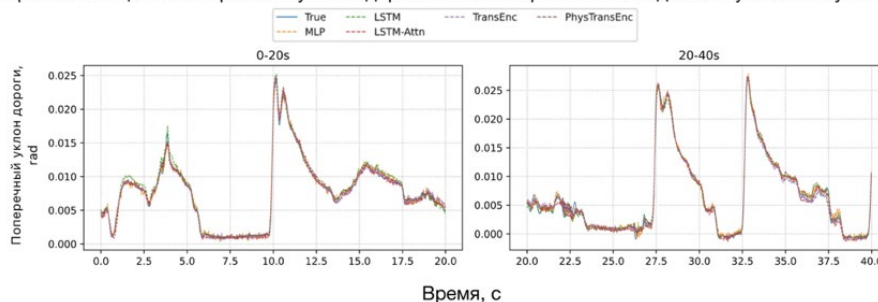


Рис. 9. – Сравнение оценки поперечного уклона дороги на основе различных моделей глубокого обучения.

Сравнение (рис. 10) показывает, что варианты А и В точнее отслеживают эталонные кривые, однако более чувствительны к измерительному шуму. Вариант С формирует более сглаженную оценку с небольшой задержкой, подавляя высокочастотные колебания и тем самым повышая устойчивость результата. По функции накопленного распределения ошибок (рис. 11а) видно, что для вариантов А и В в большинстве случаев модуль ошибки ниже 0,01 рад, тогда как для варианта С характерный уровень составляет около 0,015 рад. Точечные диаграммы (рис. 11b) подтверждают меньший разброс у варианта А; при этом ошибки варианта С остаются в пределах 0,02 рад. Среднеквадратичная ошибка равна 0,008 рад (вариант А), 0,010 рад (вариант В) и 0,013 рад (вариант С). Несмотря на меньшую точность, вариант С обеспечивает более устойчивую и согласованную оценку параметров движения.

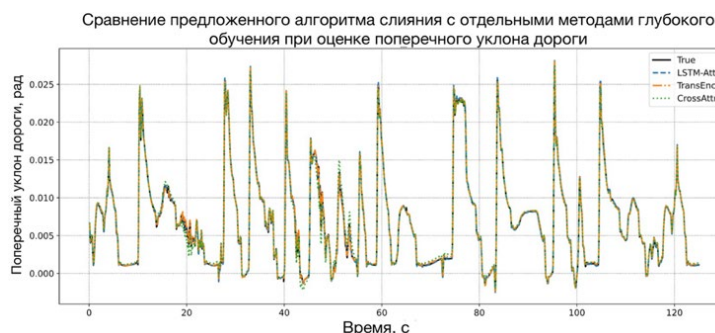


Рис. 10. – Сравнение предложенного алгоритма слияния и отдельных методов глубокого обучения.

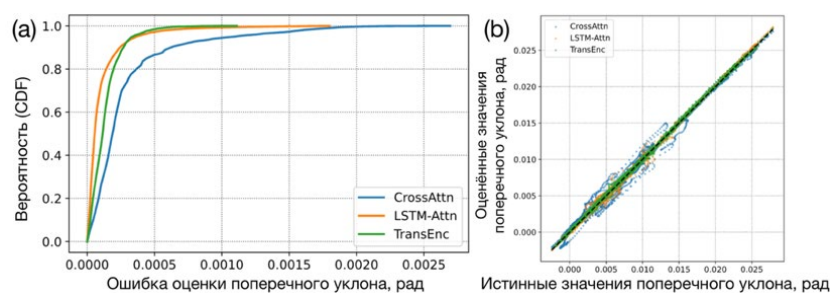


Рис. 2. – (a) Сравнение функций накопленного распределения (CDF) для оценки поперечного уклона дороги разными алгоритмами; (b) сравнение по точечным диаграммам для тех же алгоритмов.

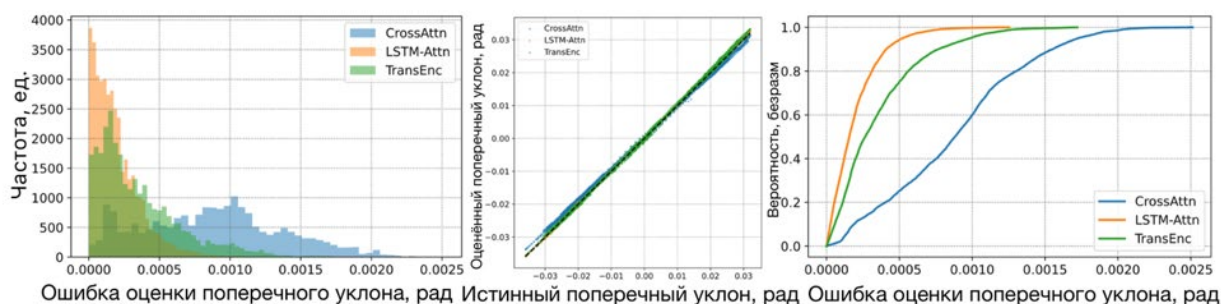


Рис. 3. – Поперечный уклон: гистограммы, scatter, CDF по алгоритмам.

Помимо среднеквадратичной ошибки для движущихся объектов принципиально важна оценка в «худшем случае», поскольку редкие пиковые отклонения могут приводить к критическим режимам движения. В качестве такой оценки используется максимальная абсолютная ошибка, то есть наибольшее по времени абсолютное отклонение оценки от эталонного значения на интервале наблюдения $[0, T]$. Этот показатель характеризует наиболее неблагоприятный момент, тогда как среднеквадратичная ошибка отражает усреднённое качество.

Для оценки массы автомобиля (при исключении интервала начальной сходимости в первые две секунды) максимальная абсолютная ошибка в переходных режимах достигает порядка 200 кг, тогда как в стационарных режимах отклонение удерживается в пределах 50 кг. Для продольного уклона наблюдаются пиковые отклонения до 0,10 рад, а при исключении начального участка характерный максимум снижается примерно до 0,06 рад. Для

поперечного уклона среднеквадратичная ошибка составляет 0,008 рад (вариант А), 0,010 рад (вариант В) и 0,013 рад (вариант С), а максимальная абсолютная ошибка равна 0,015 рад, 0,017 рад и 0,020 рад соответственно. Наибольшие отклонения проявляются в переходных режимах, поэтому дальнейшая работа будет направлена на их снижение за счёт повышения робастности в динамических режимах, адаптивной настройки ковариаций шумов и расширения обучающей выборки сценариями с резкими изменениями ускорений и уклонов.

Предложенная объединяющая схема динамически перераспределяет веса между нейросетевой моделью и алгоритмом SR-UKF (рис. 13). Вес, соответствующий оценке массы, в большинстве режимов близок к единице, что указывает на доминирование нейросетевой оценки. Для продольного и поперечного уклонов веса, как правило, удерживаются на уровне 0,7–0,9. Резкие изменения весов совпадают с переходными режимами движения и отражают смену динамики объекта управления.

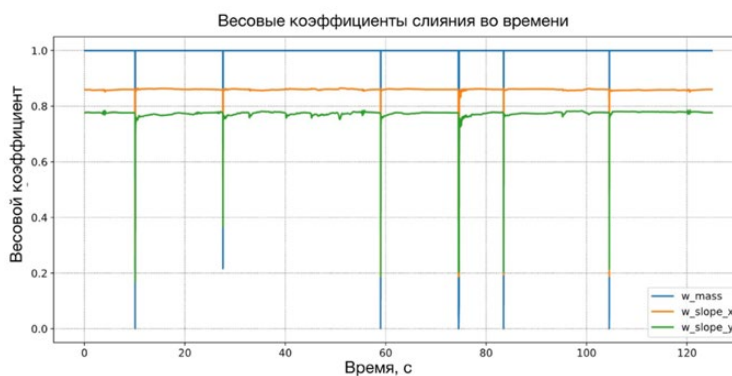


Рис. 4. – Веса слияния для каждой переменной во времени.

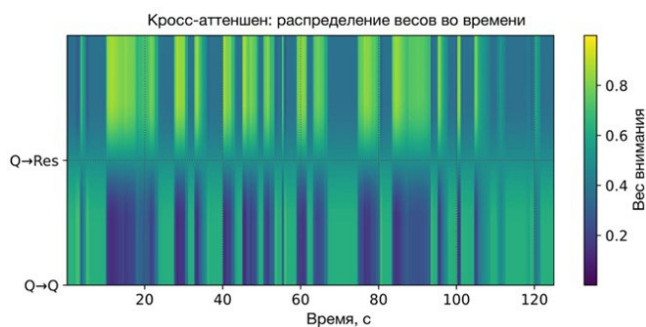


Рис. 5. – Тепловая карта распределения внимания во времени.

Тепловая карта распределения внимания предложенной модели приведена на рис. 14. По вертикальной оси показано распределение внимания между различными компонентами входной информации, по горизонтальной — время. Модель устойчиво сохраняет вклад остаточных компонент, а всплески внимания совпадают с динамическими переходами. Это означает, что сеть выделяет критические участки траектории, на которых изменяются параметры движения (например, при резких изменениях крена или ускорений), и перераспределяет веса в пользу наиболее информативных признаков. Таким образом, механизм внимания повышает интерпретируемость процесса объединения оценок и способствует более точному восстановлению параметров (уклоны, масса, ускорения) именно в переходных режимах, где традиционные методы демонстрируют наибольшие ошибки.

Заключение

Представлена гибридная схема оценки состояния автомобиля, объединяющая адаптивный алгоритм SR-UKF и нейросетевую модель с физическими ограничениями на основе энкодера трансформера. Предложенный подход обеспечивает высокоточную оценку массы и уклонов дороги, сочетая интерпретируемость фильтрационных методов и гибкость нейросетевого моделирования как в стационарных, так и в динамических режимах движения.

Для компенсации запаздывания SR-UKF и снижения чувствительности нейросетевой оценки к шуму используется механизм внимания, который объединяет выходы обоих оценщиков и обучает временные веса их вклада. Это повышает точность и робастность оценивания параметров движения транспортного средства. Для ресурсно-ограниченных платформ применяется перенос знаний: поведение крупной модели-преподавателя переносится в компактную рекуррентную модель типа долговременной краткосрочной

памяти, что заметно снижает вычислительные затраты при минимальной потере точности. Несмотря на несколько худшую точность в переходных режимах по сравнению с чисто нейросетевыми решениями, преимущества по интерпретируемости, устойчивости и пригодности к встраиванию делают предложенный подход практичным для реальных автомобильных систем.

В дальнейшем планируется расширение входных данных за счёт лидарных и видеодатчиков для прогнозирования изменений дорожных условий и последующей совместной оптимизации управления транспортным средством с целью повышения безопасности и комфорта пассажиров [19].

Литература (References)

1. Guo J., He C., Li J., Wei H. Energies. 2022. Vol. 15. № 11. P. 4126. DOI: 10.3390/en15114126.
2. Liao X., Hu Q., Sun D., Liu W., Han W. Real-time Road slope estimation based on adaptive extended Kalman filter algorithm with in-vehicle data. Proc. Chinese Control and Decision Conf. (CCDC). Chongqing. 2017. Pp. 6889–6894. DOI: 10.1109/CCDC.2017.7978422.
3. Palella N., Colombo L., Pisoni F., Avellone G., Philippe V. Sensor fusion for land vehicle slope estimation. Proc. DGON Inertial Sensors and Systems (ISS). Karlsruhe. 2016. Pp. 1–20. DOI: 10.1109/InertialSensors.2016.7745683.
4. Johansson K. Road Slope Estimation with Standard Truck Sensors. Stockholm. KTH Royal Institute of Technology. 2005. 65 p.
5. Klomp M., Gao Y., Bruzelius F. Vehicle System Dynamics. 2014. Vol. 52. Suppl. 1. Pp. 172–188. DOI: 10.1080/00423114.2014.887737.
6. Mahyuddin M.N., Na J., Herrmann G., Ren X., Barber P. IEEE Trans. Ind. Electron. 2014. Vol. 61. № 6. Pp. 2851–2863. DOI: 10.1109/TIE.2013.2276020.
7. Li B., Zhang J., Du H., Li W. Measurement. 2017. Vol. 95. Pp. 439–455. DOI: 10.1016/j.measurement.2016.10.045.

8. Zhao M., Yang F., Sun D., Han W., Xie F., Chen T. A joint dynamic estimation algorithm of vehicle mass and road slope considering braking and turning. Proc. Chinese Control and Decision Conf. (CCDC). Shenyang. 2018. Pp. 5868–5873. DOI: 10.1109/CCDC.2018.8408157.
 9. Liao Y., Hu Z., Jia H., Tian Y., Zhong S., Peng X. Automotive Engineer. 2024. № 10. Pp. 1–7. DOI: 10.20104/j.cnki.1674-6546.2024270.
 10. Feng J., Qin D., Liu Y., et al. Measurement. 2021. Vol. 181. P. 109609. DOI: 10.1016/j.measurement.2021.109609.
 11. Qin T., Yao Z., Fan H., Xia R., Chen T. IFAC-PapersOnLine. 2023. Vol. 56. № 2. Pp. 3066–3071. DOI: 10.1016/j.ifacol.2023.10.1436.
 12. Yan Y., Li H. Sensors. 2022. Vol. 22. № 5. P. 1867. DOI: 10.3390/s22051867.
 13. Ustunel E., Masazade E. IET Intell. Transp. Syst. 2019. Vol. 13. № 10. Pp. 1590–1602. DOI: 10.1049/iet-its.2018.5479.
 14. Cohen N., Klein I. Eng. Appl. Artif. Intell. 2025. Vol. 146. P. 110221. DOI: 10.1016/j.engappai.2025.110221.
 15. Rajamani R. Vehicle Dynamics and Control. Boston. Springer US. 2006. 470 p.
 16. Hou T., Kulik A.A. Tire—Road Friction Characteristics Estimation Based on ESO Observer and LSTM Neural Networks. Proc. 2025 Int. Russian Smart Industry Conf. (SmartIndustryCon). Sochi. 2025. Pp. 946–952. DOI: 10.1109/SmartIndustryCon65166.2025.10986274.
 17. Vaswani A., Shazeer N., Parmar N., et al. Attention Is All You Need. Advances in Neural Information Processing Systems. 2017. Pp. 5998–6008.
 18. Yaskelyainen S.D., Rego G.E., Volkova E.M., Kirillov I.S., Krasnyy V.A., Melnikov D.S. Inzhenernyj vestnik Dona. 2025. №3. URL: ivdon.ru/ru/magazine/archive/n3y2025/9892.
-



19. Kozyrev S.V., Polyantseva K.A. Inzhenernyj vestnik Dona. 2024. № 11. URL: ivdon.ru/ru/magazine/archive/n11y2024/9625.

Дата поступления: 16.12.2025

Дата публикации: 6.02.2026