Сайпулаев Г.Р., Половинкин Р.А., Сайпулаева Г.А. Построение системы оптимального управления движением меканум-платформы методом «актер-критик»// Транспортное, горное и строительное машиностроение: наука и производство. – 2026. – № 40. – С. 18-24.

УДК 531.133.1 https://doi.org/10.26160/2658-3305-2026-40-18-24

ПОСТРОЕНИЕ СИСТЕМЫ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ ДВИЖЕНИЕМ МЕКАНУМ-ПЛАТФОРМЫ МЕТОДОМ «АКТЕР-КРИТИК»

Сайпулаев Г.Р., Половинкин Р.А., Сайпулаева Г.А.
Национальный исследовательский университет «МЭИ», Москва

Ключевые слова: меканум-платформа, метод «актер-критик», динамика, стабилизация движения, условие непроскальзывания, нейронная сеть, обучение с подкреплением, оптимальное управление.
Аннотация. В рамках данной работы ставится задача построения алгоритма управления методом «актер-критик» для модели динамики меканум-платформы, в котором оптимизируются энергозатраты на управление и отклонения от программного движения. Для этого были построена стационарная система нелинейных дифференциальных уравнений динамики меканум-платформы относительно расширенного вектора состояния. Сформулирована задача оптимального управления и предложено её приближенное решение в виде системы управления, использующей нейронные сети актера и критика, настройка весов которых проводится методом обучения с подкреплением. Полученное математическое описание системы управления меканум-платформы может быть использовано для реализации заданных программных движений. По результатам численного моделирования продемонстрирована работоспособность предложенного управления и сделан вывод о том, что после окончания переходных процессов предложенное управление устраняет отклонения скоростей платформы от программного движения, при этом поправки к управляющим воздействиям обнуляются. Предложен вариант базисных функций для нейронных сетей актера и критика, с помощью которых обеспечивается достаточно точное управление нелинейной динамикой меканум-платформы. Результаты работы могут быть использованы при построении прототипов и разработке программного обеспечения колесных роботов.

DESIGNING OPTIMAL MOTION CONTROL SYSTEM FOR A MECANUM-PLATFORM USING THE ACTOR-CRITIC METHOD

Saypulaev G.R., Polovinkin R.A., Saypulaeva G.A.
National Research University “Moscow Power Engineering Institute”, Moscow

Keywords: mecanum-platform, actor-critic method, dynamics, motion stabilization, no-slip condition, neural network, reinforcement learning, optimal control.
Abstract. This study aims to develop an Actor-Critic control algorithm for the dynamic model of a Mecanum-platform, optimizing both control energy consumption and deviations from the reference trajectory. To achieve this, a stationary system of nonlinear differential equations describing the platform's dynamics was formulated in terms of an extended state vector. An optimal control problem is formulated, and an approximate solution is proposed in the form of a control system employing actor-critic neural networks, the weights of which are tuned using reinforcement learning. The resulting mathematical description of the control system can be used to implement desired reference trajectories. Numerical simulations demonstrate the effectiveness of the proposed control scheme and lead to the conclusion that, following the completion of transient processes, the proposed control eliminates deviations of the platform's velocities from the desired motion, while the corrections to the control inputs become equal to zero. A set of basis functions for the actor and critic neural networks is proposed, enabling sufficiently accurate control of the platform's nonlinear dynamics. The results of the work can be applied to the prototyping and software development of wheeled robots.

Введение. В связи с растущими требованиями к разработкам высокоточных мобильных роботов, применяемых в различных сферах человеческой деятельности, возникает потребность разрабатывать системы оптимального управления движением на основе нелинейных моделей движения [1, 2]. Обзор [1] посвящен анализу методов оптимального и автономного управления на основе обучения с подкреплением для дискретных и непрерывных систем. Подобные методы позволят компенсировать погрешности позиционирования мобильных роботов, вызванных различными факторами [3].

В качестве одного из перспективных методов оптимального управления нелинейными системами отмечается [2], так называемый, метод «актер-критик» (Actor-Critic). Суть метода заключается в алгоритме одновременной аппроксимации целевых функционалов,

оценивающих качество реализации движения (роль критика) и оптимизирующих цель управления (роль актера).

В работе [4] задача оптимального управления, формулируется как минимизация интегрального функционала, характеризующего затраты на управление и отклонения от программного движения. При этом решение уравнения Гамильтона–Якоби–Беллмана в частных производных аппроксимируется с помощью нейронных сетей критика и актера. Для обновления весов нейронной сети критика используется метод наименьших квадратов, а для нейронной сети актера – метод градиентный метод наименьших квадратов. Для придания адаптивности методу управления используется метод конкурентного обучения, оценивающий неизвестных параметров робота. Представлено компьютерное моделирование работы алгоритма на примере кинематическое модели робота с дифференциальным приводом, подтверждающее устойчивость системы и сходимость алгоритма.

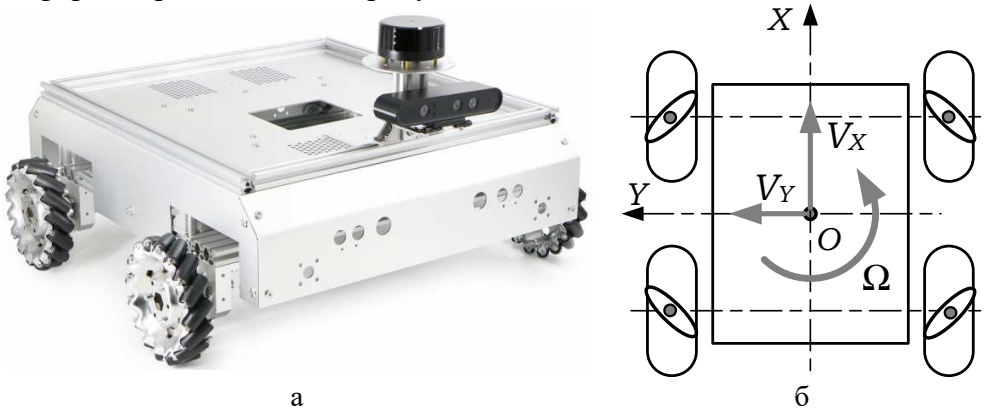
Применение управления методом «актер-критик» для непрерывных по времени систем демонстрируется в задаче следования по траектории для динамической модели робота с дифференциальным приводом [5] и динамической модели платформы при учете упрощенной модели меканум-колес и отсутствии проскальзывания точек контакта [6].

Статья [7] посвящена получению оптимальных решений для управления нелинейными системами с дискретным временем. Предлагается рекурсивный метод наименьших квадратов для последовательного приближения решения обобщенного уравнения Гамильтона–Якоби–Беллмана. Для аппроксимации и настройки закона управления используется нейронная сеть. Численные эксперименты на модели двухзвенного манипулятора подтверждает асимптотическую стабильность и сходимость к оптимальному управлению.

В работе [8] управление методом «актер-критик» используется для построения субоптимального управления, выводящего систему на скользящий режим. А в статье [9] управление методом «актер-критик» предлагается алгоритм управления методом «актер-критик» при наличии ограничений на управляющие воздействия.

В данной работе ставится задача построения алгоритма управления методом «актер-критик» для модели динамики меканум-платформы, в котором оптимизируются энергозатраты на управление и отклонения от программного движения. При этом само программное движение является заданным строго и не подлежит оптимизации. Математическая модель динамики робота с учетом сил линейного трения в механических передачах колес была представлена в статье [10].

Материалы и методы исследований. Рассматривается мобильная платформа с меканум-колесами, пример которой представлен на рисунке 1, а. Кинематическая схема меканум-платформы представлена на рисунке 1, б.



а б

Рис. 1. Мобильная платформа с меканум-колесами: а) пример робота Mecabot; б) кинематическая схема

В качестве уравнений динамики меканум-платформы, движущейся по горизонтальной плоскости, при учете условий непроскальзывания точек контакта с опорной поверхностью примем модель [10]:

$$
\begin{aligned}
&m_a\dot{V}_X - m_s\Omega V_Y + \zeta_X V_X = F_X,\\
&m_a\dot{V}_Y + m_s\Omega V_X + \zeta_Y V_Y = F_Y,\\
&I_a\dot{\Omega} + \mu_\Omega\Omega = M_\Omega,
\end{aligned}
\tag{1}
$$

где m_a, I_a – приведенные масса и момент инерции системы; m_s – масса робота, состоящего из платформы и четырех колес; $\zeta_X, \zeta_Y, \mu_\Omega$ – коэффициенты линейного трения в механических передачах колес; F_X, F_Y, M_Ω – управляющие воздействия, связанные с моментами двигателей колес; V_X, V_Y – проекции линейной скорости геометрического центра *O* платформы на продольную *X* и поперечную *Y* оси симметрии; Ω – угловая скорость вращения платформы вокруг вертикальной оси симметрии. Точкой над переменной указана производная по времени.

Разрешив систему (1) относительно производных, получим:

$$
\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{u}, \tag{2}
$$

где $\mathbf{x} = (V_X \quad V_Y \quad \Omega)^{\mathrm{T}}$ – вектор переменных состояния; $\mathbf{u} = (u_X \quad u_Y \quad u_\Omega)^{\mathrm{T}}$ – вектор управлений; $\mathbf{f}(\mathbf{x}) = (-\beta_X V_X + \alpha\Omega V_Y \quad -\beta_Y V_Y - \alpha\Omega V_X \quad -\beta_Z\Omega)^{\mathrm{T}}$ – вектор функций, характеризующих свободные (баллистические) движения системы; $\beta_X = \zeta_X / m_a$, $\beta_Y = \zeta_Y / m_a$, $\beta_Z = \mu_\Omega / I_a$ – коэффициенты, характеризующие линейное трение; $u_X = F_X / m_a$, $u_Y = F_Y / m_a$, $u_\Omega = M_Z / I_a$ – приведенные управления; $\alpha = m_s / m_a < 1$ – безразмерный коэффициент. Здесь и далее символ «Т» в верхнем индексе означает операцию транспонирования вектора или матрицы.

Задавая программное движение робота $\mathbf{x} = \mathbf{x}^d(t)$, как функцию времени *t*, имеем уравнения (2) на программном движении:

$$
\dot{\mathbf{x}}^d = \mathbf{f}(\mathbf{x}^d) + \mathbf{u}^d, \tag{3}
$$

откуда можно определить управление на программном движении $\mathbf{u}^d = \dot{\mathbf{x}}^d - \mathbf{f}(\mathbf{x}^d)$ и получить уравнения в отклонениях $\Delta\mathbf{x} = \mathbf{x} - \mathbf{x}^d$ от программного движения в виде:

$$
\Delta\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}^d, \Delta\mathbf{x}) + \Delta\mathbf{u}, \tag{4}
$$

где $\Delta\mathbf{u} = \mathbf{u} - \mathbf{u}^d$ – поправка к управлению; а вектор-функция $\mathbf{F}(\mathbf{x}^d, \Delta\mathbf{x}) = \mathbf{f}(\mathbf{x}^d + \Delta\mathbf{x}) - \mathbf{f}(\mathbf{x}^d)$ определяется выражением:

$$
\mathbf{F}\left(\mathbf{x}^d, \Delta\mathbf{x}\right) = \begin{pmatrix} -\beta_X\Delta V_X + \alpha\left(V_Y^d\Delta\Omega + \Omega^d\Delta V_Y + \Delta\Omega\Delta V_Y\right) \\ -\beta_Y\Delta V_Y - \alpha\left(V_X^d\Delta\Omega + \Omega^d\Delta V_X + \Delta\Omega\Delta V_X\right) \\ -\beta_Z\Delta\Omega \end{pmatrix}. \tag{5}
$$

Для приведения уравнений динамики к уравнениям автономной системы введем расширенные векторы состояния $\mathbf{z} = (\Delta\mathbf{x} \quad \mathbf{x}^d)^{\mathrm{T}}$ и управления $\mathbf{u}_z = (\Delta\mathbf{u} \quad \mathbf{u}^d)^{\mathrm{T}}$. Тогда получим уравнения (3) и (4) в виде:

$$
\dot{\mathbf{z}} = \mathbf{F}_z(\mathbf{z}) + \mathbf{u}_z, \tag{6}
$$

где $\mathbf{F}_z\left(\mathbf{z}\right) = \left(\mathbf{F}(\mathbf{x}^d, \Delta\mathbf{x}) \quad \mathbf{f}(\mathbf{x}^d)\right)^{\mathrm{T}}$.

В качестве цели управления выберем минимизацию функционала $V(\mathbf{z}, \mathbf{u}_z)$, характеризующего энергозатраты и отклонения от программного движения:

$$
V\left(\mathbf{z}, \mathbf{u}_z\right) = \int_0^\infty \left(\mathbf{z}^{\mathrm{T}}\mathbf{Q}\mathbf{z} + \mathbf{u}_z^{\mathrm{T}}\mathbf{R}\mathbf{u}_z\right) dt \to \min, \tag{7}
$$

где $\mathbf{Q} = \mathrm{diag}(q_1, q_2, q_3, 0, 0, 0), \mathbf{R} = \mathrm{diag}(r_1, r_2, r_3, 0, 0, 0)$ – диагональные матрицы $(q_i > 0, r_i > 0, i = 1, 2, 3)$, характеризующие штрафы за отклонения от программного движения и энергозатраты на управление.

Введем функцию Гамильтона:

$$H = \mathbf{z}^{\mathrm{T}}\mathbf{Q}\mathbf{z} + \mathbf{u}_z^{\mathrm{T}}\mathbf{R}\mathbf{u}_z + \left(\frac{\partial V}{\partial \mathbf{z}}\right)^{\mathrm{T}}\left(\mathbf{F}(\mathbf{z}) + \mathbf{u}_z\right), \tag{8}$$

тогда уравнение для функции $V(\mathbf{z}, \mathbf{u}_z)$ имеет вид уравнения Гамильтона–Якоби–Беллмана:

$$\frac{\partial V}{\partial t} + \min\left[H\left(\mathbf{z}, \mathbf{u}_z, \frac{\partial V}{\partial \mathbf{z}}\right)\right] = 0. \tag{9}$$

Для оптимального значения $V^*(\mathbf{z}^*, \mathbf{u}_z^*)$ выполняется:

$$\min\left[H\left(\mathbf{z}, \mathbf{u}_z, \frac{\partial V^*}{\partial \mathbf{z}}\right)\right] = 0, \tag{10}$$

а из необходимого условия минимума по управлению имеем:

$$\Delta\mathbf{u}^* = -\frac{1}{2}\mathrm{diag}\left(r_1^{-1}, r_2^{-1}, r_3^{-1}\right)\left(\frac{\partial V^*}{\partial \Delta\mathbf{x}}\right)^{\mathrm{T}}, \quad \mathbf{u}^d = \dot{\mathbf{x}}^d - \mathbf{f}(\mathbf{x}^d). \tag{11}$$

Здесь учтено, что $V(\mathbf{z}, \mathbf{u}_z)$ не зависит от $\mathbf{x}^d, \mathbf{u}^d$ в силу заданного вида матриц $\mathbf{Q}, \mathbf{R}$. Звездочка в верхних индексах переменных означает оптимальную зависимость величины.

Следуя методу «актер-критик» используем аппроксимацию функции Беллмана однослойной нейронной сетью критика:

$$V\left(\mathbf{z}(t)\right) = \mathbf{W}_C^{\mathrm{T}}\boldsymbol{\varphi}\left(\mathbf{z}(t)\right) + \varepsilon_B, \tag{12}$$

где $\mathbf{W}_C$ – столбец весов критика размерности N_φ; $\boldsymbol{\varphi}(\mathbf{z})$ – столбец базисных функций нейронной сети размерности N_φ, а ошибка аппроксимации ε_B имеет вид:

$$\varepsilon_B = \int_{t-T_*}^{t}\left(\mathbf{z}^{\mathrm{T}}\mathbf{Q}\mathbf{z} + \mathbf{u}_z^{\mathrm{T}}\mathbf{R}\mathbf{u}_z\right)dt + \mathbf{W}_C^{\mathrm{T}}\left[\boldsymbol{\varphi}\left(\mathbf{z}(t)\right) - \boldsymbol{\varphi}\left(\mathbf{z}(t - T_*)\right)\right]. \tag{13}$$

где T_* – величина, характеризующая время обучения критика.

Выход нейронной сети критика определяется текущей матрицей весов критика $\hat{\mathbf{W}}_C$ и входом $\mathbf{z}(t)$:

$$V(\mathbf{z}) = \hat{\mathbf{W}}_C^{\mathrm{T}}\boldsymbol{\varphi}(\mathbf{z}), \tag{14}$$

а поскольку веса критика неизвестны, то закон управления (11) аппроксимируется нейронной сетью актера:

$$\Delta\mathbf{u}^* = -\frac{1}{2}\mathrm{diag}\left(r_1^{-1}, r_2^{-1}, r_3^{-1}\right)\left(\frac{\partial \boldsymbol{\varphi}(\mathbf{z})}{\partial \Delta\mathbf{x}}\right)^{\mathrm{T}}\hat{\mathbf{W}}_A, \quad \mathbf{u}^d = \dot{\mathbf{x}}^d - \mathbf{f}(\mathbf{x}^d), \tag{15}$$

где $\hat{\mathbf{W}}_A$ – текущий столбец весов актера размерности N_φ.

Для обучения нейронных сетей критика и актера используем уравнения [1], основанные на обучении с подкреплением:

$$\begin{aligned}
\frac{d\hat{\mathbf{W}}_C}{dt} &= -\alpha_1 \frac{\Delta\boldsymbol{\varphi}(\mathbf{z})}{\left(1 + \Delta\boldsymbol{\varphi}^{\mathrm{T}}(\mathbf{z})\Delta\boldsymbol{\varphi}(\mathbf{z})\right)^2}\left[\int_{t-T_*}^{t}\left(\mathbf{z}^{\mathrm{T}}\mathbf{Q}\mathbf{z} + \mathbf{u}_z^{\mathrm{T}}\mathbf{R}\mathbf{u}_z\right)dt + \Delta\boldsymbol{\varphi}^{\mathrm{T}}(\mathbf{z})\hat{\mathbf{W}}_C\right], \\
\frac{d\hat{\mathbf{W}}_A}{dt} &= -\frac{\alpha_2}{4}\left(\frac{\partial \boldsymbol{\varphi}(\mathbf{z})}{\partial \Delta\mathbf{x}}\right)\mathrm{diag}\left(r_1^{-1}, r_2^{-1}, r_3^{-1}\right)\left(\frac{\partial \boldsymbol{\varphi}(\mathbf{z})}{\partial \Delta\mathbf{x}}\right)^{\mathrm{T}}\hat{\mathbf{W}}_A\frac{\Delta\boldsymbol{\varphi}^{\mathrm{T}}(\mathbf{z})\hat{\mathbf{W}}_C}{\left(1 + \Delta\boldsymbol{\varphi}^{\mathrm{T}}(\mathbf{z})\Delta\boldsymbol{\varphi}(\mathbf{z})\right)^2} \\
&\quad - \alpha_2\left[\mathbf{A}_2\hat{\mathbf{W}}_A - \mathbf{A}_1\Delta\boldsymbol{\varphi}^{\mathrm{T}}(\mathbf{z})\hat{\mathbf{W}}_C\right],
\end{aligned} \tag{16}$$

где α_1, α_2 – параметры, характеризующие скорость обучения критика и актера; $\Delta\boldsymbol{\varphi}(\mathbf{z}) = \boldsymbol{\varphi}(\mathbf{z}(t)) - \boldsymbol{\varphi}(\mathbf{z}(t - T_*))$ – приращение вектора базисных функций; $\mathbf{A}_1$ и $\mathbf{A}_2$ – параметры

алгоритма обучения, представленные столбцом размерности N_φ и матрицей размерности $N_\varphi \times N_\varphi$.

В качестве вектора базисных функций рассмотрим

$$\boldsymbol{\varphi}(\mathbf{z}) = (\Delta V_X^2 \quad \Delta V_Y^2 \quad \Delta\Omega^2 \quad \Delta V_X \Delta V_Y \quad \Delta V_X \Delta\Omega \quad \Delta V_Y \Delta\Omega)^{\mathrm{T}}. \tag{17}$$

Полученное управление методом «актер-критик» может быть использовано для устранения отклонений от программного движения с учетом нелинейной обратной связи по скоростям платформы.

Результаты. Проведем численное моделирование движения робота. В качестве примера для численного моделирования рассмотрим параметры робота

$$\beta_X = \beta_Y = \beta_Z = 1{,}5 \text{ м/с}, \quad \alpha = 0{,}8; \tag{18}$$

и параметры системы управления:

$$T_* = 50 \text{ с}, \quad \alpha_1 = \alpha_2 = 0.2\,1/\text{с}, \quad \mathbf{A}_2 = \mathrm{diag}(1,1,1,1,1,1),$$

$$\mathbf{A}_1 = \left(1\,1/\text{с} \quad 1\,1/\text{с} \quad 1\,1/(\text{м}^2 \cdot \text{с}) \quad 1\,1/\text{с} \quad 1\,1/(\text{м} \cdot \text{с}) \quad 1\,1/(\text{м} \cdot \text{с})\right)^{\mathrm{T}}, \tag{19}$$

$$q_1 = q_2 = 1, \quad q_3 = 1\,\text{м}^2, \quad r_1 = r_2 = 1\,\text{м}^2/\text{с}^4, \quad r_3 = 1\,\text{м}^2 \cdot \text{с}^2;$$

В качестве программного движения рассмотрим

$$V_X^d = 0{,}5 \text{ м/с}, \quad V_Y^d = 0{,}7 \text{ м/с}, \quad \Omega^d = 0{,}1\,1/\text{с}; \tag{20}$$

и начальные условия

$$\begin{aligned}
&\Delta V_X(0) = 0{,}1 \text{ м/с}, \quad \Delta V_Y(0) = 0{,}2 \text{ м/с}, \quad \Delta\Omega(0) = 0{,}05\,1/\text{с},\\
&\mathrm{W}_{A1}(0) = 0{,}7 \text{ с}, \quad \mathrm{W}_{A2}(0) = 1{,}2 \text{ с}, \quad \mathrm{W}_{A3}(0) = 1{,}0 \text{ м}^2 \cdot \text{с},\\
&\mathrm{W}_{A4}(0) = 1{,}1 \text{ с}, \quad \mathrm{W}_{A5}(0) = 0{,}8 \text{ м} \cdot \text{с}, \quad \mathrm{W}_{A6}(0) = 0{,}9 \text{ м} \cdot \text{с},\\
&\mathrm{W}_{C1}(0) = 1{,}2 \text{ с}, \quad \mathrm{W}_{C2}(0) = 1{,}3 \text{ с}, \quad \mathrm{W}_{C3}(0) = 1{,}0 \text{ м}^2 \cdot \text{с},\\
&\mathrm{W}_{C4}(0) = 0{,}9 \text{ с}, \quad \mathrm{W}_{C5}(0) = 0{,}75 \text{ м} \cdot \text{с}, \quad \mathrm{W}_{C6}(0) = 0{,}85 \text{ м} \cdot \text{с},
\end{aligned} \tag{21}$$

где $\mathrm{W}_{Ai}, \mathrm{W}_{Ci}$ $(i = 1,...,6)$ – компоненты столбцов весов актера $\hat{\mathbf{W}}_A$ и критика $\hat{\mathbf{W}}_C$ соответственно.

На рисунке 2 представлены результаты моделирования стабилизации программного движения меканум-платформы.

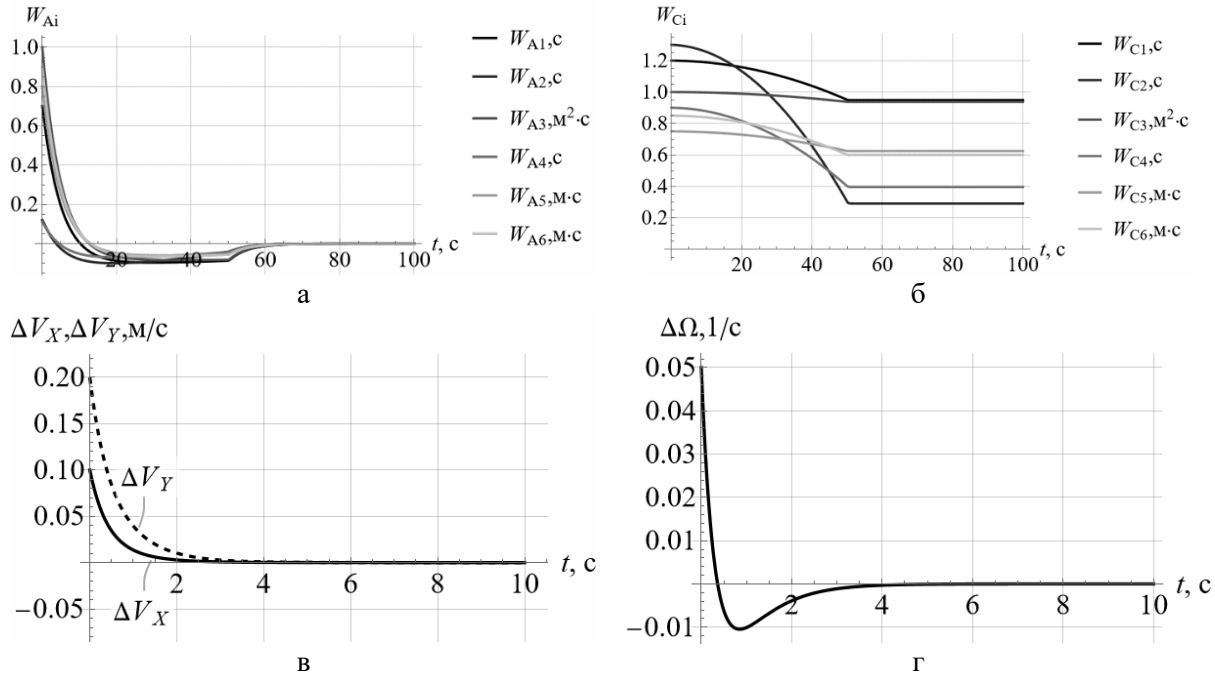


Рис. 2. Результаты моделирования: а) и б) элементы столбцов весов актера и критика; в) отклонения линейных продольной и поперечной скорости центра платформы; г) отклонение угловой скорости платформы

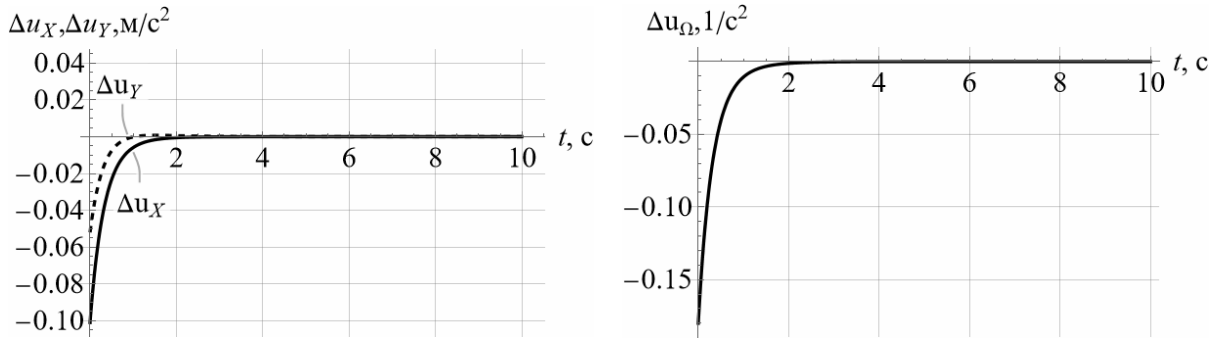


Рис. 2. Результаты моделирования: д) и е) компоненты вектора поправок к управлению

Из рисунка 2, в, г можно сделать вывод, что предложенное управление устраняет отклонения скоростей платформы от программного движения после окончания переходных процессов. При этом поправки к управляющим воздействиям после окончания переходных процессов обнуляются (рис. 2, д, е).

Отметим, что веса критика после момента времени T_* принимают постоянные значения (рис. 2, б). При этом веса актера до момента T_* выходят на установившиеся значения, а после окончания обучения нейронной сети критика веса актера экспоненциально убывают до нуля (рис. 2, а).

Выводы. По результатам численного моделирования подтверждена работоспособность предложенного закона управления движением меканум-платформы методом «Актер-Критик» с оптимизацией энергозатрат на управление и отклонений от программного движения. При этом само программное движение является заданным и остается неизменным.

Предложен вариант базисных функций для нейронных сетей актера и критика, с помощью которых обеспечивается достаточно точное управление нелинейной динамикой меканум-платформы.

По результатам работы получена математическая модель системы управления движением меканум-платформы, построенная на основе метода «Актер-Критик» и аппроксимаций нейронными сетями, использующих обучение с подкреплением. Результаты работы могут быть использованы при построении системы управления колесными аппаратами, в частности меканум-платформами.

Заключение

В данной работе с использованием метода «Актер-Критик» построено оптимальное управление движением меканум-платформы. Представлено численное моделирование управляемой динамики меканум-платформы, совершающей движение по горизонтальной плоскости. По результатам численного моделирования продемонстрирована работоспособность предложенного закона управления, оптимизирующего энергозатраты на управление и отклонения от программного движения с использованием обратной связи по скоростям.

Полученная математическая модель системы управления движением меканум-платформы может быть использована при построении прототипов и разработке программного обеспечения колесных роботов.

Список литературы

1. Kiumarsi B., Vamvoudakis K.G., Modares H., Lewis F.L. Optimal and Autonomous Control Using Reinforcement Learning: A Survey // IEEE Transactions on Neural Networks and Learning Systems. 2018, vol. 29, no. 6, pp. 2042-2062. doi.org/10.1109/TNNLS.2017.2773458.
2. Siciliano B., Khatib O. Springer Handbook of Robotics. – Berlin: Springer, 2016. – 1628 p. – doi.org/10.1007/978-3-319-32552-1.
3. Панкратьева Г.В., Мордин А.Е., Сайпулаев Г.Р. Анализ точности позиционирования при конструкционных погрешностях в установке меканум-колёс мобильной платформы // Advanced Engineering Research (Rostov-on-Don). – 2023. – Т. 23, № 4. – С. 356-364. – doi.org/10.23947/2687-1653-2023-23-4-356-364.

4-10. См. *References*.

References

1. Kiumarsi B., Vamvoudakis K.G., Modares H., Lewis F.L. Optimal and Autonomous Control Using Reinforcement Learning: A Survey // IEEE Transactions on Neural Networks and Learning Systems. 2018, vol. 29, no. 6, pp. 2042-2062. doi.org/10.1109/TNNLS.2017.2773458.
2. Siciliano B., Khatib O. Springer Handbook of Robotics. – Berlin: Springer, 2016. – 1628 p. – doi.org/10.1007/978-3-319-32552-1.
3. Pankrateva G.V., Mordin A.E., Saypulaev G.R. Analysis of Positioning Accuracy in Case of Design Errors in the Installation of Mecanum Wheels of the Mobile Platform // Advanced Engineering Research (Rostov-on-Don). 2023, vol. 23, no. 4, pp. 356-364. doi.org/10.23947/2687-1653-2023-23-4-356-364.
4. Dani A.P., Bhasin S. Adaptive Actor-Critic Based Optimal Regulation for Drift-Free Uncertain Nonlinear Systems // IEEE Open Journal of Control Systems. 2025, vol. 4, pp. 117-129. doi.org/10.1109/OJCSYS.2025.3552999.
5. Hoa D.V., Chuyen T.D., Lai L.K., Ha L.T.T. Trajectory Tracking Control for Wheeled Mobile Robot System with Uncertain Nonlinear Model based on Integral Reinforcement Learning Algorithm // International Journal of Engineering Trends and Technology. 2024, vol. 72, iss. 5. pp. 290-298. doi.org/10.14445/22315381/IJETT-V72I5P130.
6. Nguyen M.D., Ngo M.T., Quang H.D., Phuong N.D. Reinforcement Learning-Based Trajectory Control for Mecanum Robot with Mass Eccentricity Considerations // Journal of Robotics and Control. 2024, vol. 5, iss. 5, pp. 1436-1443. doi.org/10.18196/jrc.v5i5.22148.
7. Chen Z., Jagannathan S. Generalized Hamilton–Jacobi–Bellman Formulation-Based Neural Network Control of Affine Nonlinear Discrete-Time Systems // IEEE Transactions on Neural Networks. 2008, vol. 19, no. 1, pp. 90-106. doi.org/10.1109/TNN.2007.900227.
8. Fan Q.-Y., Yang G.-H. Adaptive Actor–Critic Design-Based Integral Sliding-Mode Control for Partially Unknown Nonlinear Systems with Input Disturbances // IEEE Transactions on Neural Networks and Learning Systems. 2016, vol. 27, no. 1, pp. 165-177. doi.org/10.1109/TNNLS.2015.2472974.
9. Abu-Khalaf M., Lewis F.L. Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach // Automatica. 2005, vol. 41, iss. 5, pp. 779-791. doi.org/10.1016/j.automatica.2004.11.034.
10. Adamov B.I. A Study of the Controlled Motion of a Four-wheeled Mecanum Platform // Russian Journal of Nonlinear Dynamics. 2018, vol. 14, no. 2, pp. 265-290. doi.org/10.20537/nd180209.

Сведения об авторах:	*Information about authors:*
Сайпулаев Гасан Русланович – кандидат технических наук, доцент кафедры Робототехники, мехатроники, динамики и прочности машин	**Saypulaev Gasan Ruslanovich** – candidate of technical sciences, associate professor of the Department of robotics, mechatronics, dynamics and machine strength
Половинкин Родион Александрович – студент	**Polovinkin Rodion Aleksandrovich** – student
Сайпулаева Гулизар Абдурахмановна – соискатель	**Saypulaeva Gulizar Abdurakhmanovna** – applicant
saypulaevgr@mail.ru	

Получена 20.09.2026