ADP Based Fault-Tolerant Tracking Control for Underactuated AUV with Actuators Faults via Neural Network Observer

Gaofeng Che (chexuping@163.com)
Henan Normal University  https://orcid.org/0000-0001-6620-5066

Zhen Yu
Xiamen University

Research Article

Keywords: Adaptive dynamic programming (ADP), Fault-tolerant tracking control, Actuators faults, Neural network observer, Autonomous underwater vehicle (AUV)

Posted Date: December 2nd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-807482/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
ADP based output-feedback fault-tolerant tracking control for underactuated AUV with actuators faults via neural network observer

Gaofeng Che · Zhen Yu*

Received: date / Accepted: date

Abstract In this work, the fault-tolerant tracking control issue of underactuated autonomous underwater vehicle (AUV) with actuators faults is investigated. Firstly, an output-feedback error tracking system is constructed based on the theoretical model of underactuated AUV with actuators faults. Then, an adaptive dynamic programming (ADP) based fault-tolerant control controller is developed. In our proposed control scheme, a neural-network observer is designed to approximate the system states with actuators faults. A novel ADP scheme is constructed with critic neural network and action neural network in order to reduce the jitter in the control input and improve the tracking accuracy. Based on Lyapunov approach, the stability of the error tracking system is guaranteed by the proposed controller. At last, the simulation results show that the underactuated AUV achieves better tracking performance.

Keywords Adaptive dynamic programming (ADP) · Fault-tolerant tracking control · Actuators faults · Neural network observer · Autonomous underwater vehicle (AUV)

1 Introduction

Trajectory tracking is a complex motion control task for autonomous underwater vehicle (AUV) in an unknown underwater environment (Che et al (2019a); Qiao and Zhang (2017); Shen et al (2018); Che et al (2019b)). Many trajectory-tracking control methods have been developed for AUV without actuators faults, such as adaptive terminal-sliding-mode control method (Qiao and Zhang (2017); Zhang et al (2018a)), fuzzy control method (Liu et al (2019));...

Actuators are very important parts of underactuated AUV. The actuators faults may lead to performance degradation of underactuated AUV (Hao et al (2019); Kadiyam et al (2020)). In order to maintain system stability and the acceptable tracking accuracy, many fault-tolerant control strategies have been developed for AUV with actuators faults, such as adaptive fault-tolerant control method (Liu et al (2018a)), adaptive terminal sliding mode based fault-tolerant control method (Zhang et al (2015)), backstepping based adaptive region-tracking fault-tolerant control method (Zhang et al (2017)) and so on.

The adaptive dynamic programming (ADP) is introduced into this work to transform the trajectory-tracking control problem into optimal control problem for underactuated AUV with actuators faults. Policy iteration (PI) algorithm and value iteration (VI) algorithm are two important ADP procedures to solve the complex Hamilton-Jacobi-Bellman (HJB) equation (Gong et al (2019); Liu et al (2018b); Sun and Liu (2018)).

Many ADP algorithms have been developed to solve the tracking control problems for nonlinear systems in recent years. An infinite-time optimal tracking control problem is investigated based on greedy heuristic dynamic programming (HDP) iteration algorithm (Zhang et al (2008)). The output tracking control problem is solved based on event-driven ADP scheme (Zhang et al (2018b)). The time delays are considered and HDP is designed to solve the tracing control problem for a class of nonlinear systems (Zhang et al (2011)). The ADP based tracking control scheme is designed for coal gasification system (Zhang et al (2014)). The ADP algorithm is designed for tracking control with unknown system dynamics (Kiumarisi and Lewis (2015); Qin et al (2014)). The tracking controller based on ADP scheme is designed for fully-actuated AUV with current disturbances and rudders faults. The neural-network estimators are employed to approximate the current disturbances and rudder faults (Che and Yu (2020)). A PI algorithm is developed for online fault compensation control of a class of affine nonlinear systems with actuators failures (Zhao et al (2016)).

The main contribution of this work can be summarized as follows:

- An output-feedback error tracking system is constructed based on the theoretical model of underactuated AUV with actuators faults.
- The neural network observer is designed to approximate the actuators faults. The approximate actuators faults is introduced into an improved performance index function based on the performance index (Zhang et al (2008); Wei et al (2018); Zhao et al (2017)).
- Because AUV is a very complex nonlinear system, the critic-action neural networks are employed in order to reduce jitter in the control input which are different from the single critic neural network structure(Zhao et al (2017)).
The error tracking system can be guaranteed to be uniformly ultimately bounded (UUB) based on the Lyapunov stability theorem.

The rest of the paper is organized as follows. The output feedback based error tracking system is constructed and problem formulation is described in Section 2. In Section 3, the fault-tolerant ADP tracking controller with neural network observer is designed. Simulation examples are provided to demonstrate the effectiveness of the proposed method in Section 4. The conclusion is drawn in Section 5.

2 Theoretical model of underactuated AUV and problem formulation

2.1 Theoretical model of underactuated AUV

Two coordinate systems are used in the theoretical model of underactuated AUV as shown in Fig. 1. One is the inertial coordinate system \( \{O_e, X_e, Y_e, Z_e\} \) and the other is the body-fixed coordinate system \( \{O_b, X_b, Y_b, Z_b\} \).

The theoretical model of underactuated AUV without actuators faults is presented as follows:

\[
\begin{aligned}
\dot{\eta} &= J(\eta)\xi \\
M\ddot{\xi} + C(\xi)\xi + D(\xi)\dot{\xi} + g(\eta) &= \tau
\end{aligned}
\]  

(1)

where \( \eta = [x\ y\ z\ \phi\ \theta\ \psi]^T \) is the location vector with respect to the inertial coordinate system; \( \xi = [u\ v\ w\ p\ q\ r]^T \) is the velocity vector with respect to the body-fixed coordinate system; \( M \in \mathbb{R}^{6\times6} \) is the inertia matrix; \( C(\xi) \in \mathbb{R}^{6\times6} \) is the Coriolis and centripetal matrix; \( D(\xi) \in \mathbb{R}^{6\times6} \) is the hydrodynamic damping matrix; \( g(\eta) \in \mathbb{R}^{6\times1} \) is the gravitational forces and moment vector; \( \tau \in \mathbb{R}^{6\times1} \) is the control forces; \( J(\eta) \in \mathbb{R}^{6\times6} \) is the spatial transformation matrix between two coordinate systems.
The kinematics of underactuated AUV is described as follows:
\[
\begin{align*}
\xi &= J^{-1}\dot{\eta} \\
\dot{\xi} &= J^{-1}\ddot{\eta} - J^{-1}JJ^{-1}\eta
\end{align*}
\] (2)

where \( J = J(\eta) \).

Combining equation (1) with equation (2), we can get
\[
\ddot{\eta} = (MJ^{-1})^{-1}(MJ^{-1}\dot{J}J^{-1} - CJ^{-1} - DJ^{-1})\dot{\eta} - (MJ^{-1})^{-1}g + (MJ^{-1})^{-1}\tau
\] (3)

where \( C = C(\xi) \); \( D = D(\xi) \); \( g = g(\eta) \).

2.2 Problem formulation

The desired trajectory is given as follows:
\[
\ddot{\eta}_d = (M_dJ_d^{-1})^{-1}(M_dJ_d^{-1}\dot{J}J^{-1} - C_dJ_d^{-1} - D_dJ_d^{-1})\dot{\eta}_d - (M_dJ_d^{-1})^{-1}\dot{g}_d
\] (4)

The error vectors are defined as follows:
\[
\begin{align*}
e_{\eta} &= \eta - \eta_d \\
e_{\tau} &= \tau - \tau_d
\end{align*}
\] (5)

Then substituting equations (4), (5) into equation (3), the output feedback based error tracking system is given as follows:
\[
\dot{e}_{\eta} = (MJ^{-1})^{-1}(MJ^{-1}\dot{J}J^{-1} - C_dJ_d^{-1} - D_dJ_d^{-1})e_{\eta} + \Theta + (MJ^{-1})^{-1}e_{\tau}
\] (6)

where \( \Theta = (MJ^{-1})^{-1}(M_dJ_d^{-1} - MJ^{-1})\dot{\eta}_d + (MJ^{-1})^{-1}(MJ^{-1}\dot{J}J^{-1} - CJ^{-1} - DJ^{-1} - M_dJ_d^{-1}\dot{J}J_d^{-1} + C_dJ_d^{-1} + D_dJ_d^{-1})\dot{g}_d + (MJ^{-1})^{-1}(\dot{g}_d - g) \).

We define error vector \( x = [e_{\eta} \ e_{\tau}]^T \), then the error tracking system (6) can be transformed as follows:
\[
\dot{x} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} I \\ (MJ^{-1})^{-1} \end{bmatrix} x + \begin{bmatrix} 0 \\ \Theta \end{bmatrix} e_{\tau}
\] (7)

where \( I \in \mathbb{R}^{6 \times 6} \) is the identity matrix.

Actuators faults are described as \( f \in \mathbb{R}^{m \times 1} \) and \( m \) is the number of actuators. The real output of actuators with actuators faults is given as follows:
\[
\mu' = \mu - f
\] (8)

where \( \mu \in \mathbb{R}^{m \times 1} \) is the output of controller.

The vector of control forces and control torque \( e'_{\tau} \) of underactuated AUV with actuators faults can be represented as follows:
\[
e'_{\tau} = e_{\tau} - \tau_f = B\mu' = B\mu - Bf
\] (9)
where $B \in \mathbb{R}^{6 \times m}$ is the actuators configuration matrix; $e_{r} = B \mu$; $\tau_{f} = Bf$.

The error tracking system with actuators faults is given as follows:

$$
\dot{x} = \varpi(x) + \rho(x)(\mu - f)
$$

(10)

where $\varpi(x) = \begin{bmatrix}
0 & I \\
0 & (MJ^{-1})^{-1}(MJ^{-1}JJ^{-1} - CJ^{-1} - DJ^{-1})
\end{bmatrix} x + \begin{bmatrix}
0 \\
\Theta
\end{bmatrix}$, $\rho(x) = \begin{bmatrix}
0 \\
(MJ^{-1})^{-1}B
\end{bmatrix}$.

**Assumption 1** Because underactuated AUV does not have independent actuators in the sway and heave axes, the available controls are the surge force, pitch moment and the yaw moment. The actuators faults $f$ satisfies that $\|f\| = \|K\mu\| \leq \|\mu\| \leq \delta_{1}$. $K$ is a diagonal matrix and element $k_{ii}$ of diagonal matrix $K$ satisfies $0 \leq k_{ii} < 1$. $\varpi(x)$ and $\rho(x)$ are locally Lipchitz continuous. $\delta_{1}$ is a positive constant.

The performance index function is defined as follows:

$$
V_{1}(x, \mu) = \int_{t}^{\infty} e^{\gamma(t-\sigma)}(\beta \tilde{f}^{T}(\sigma) \tilde{f}(\sigma) + U(x(\sigma), \mu(\sigma)))d\sigma
$$

(11)

where $U(x, \mu) = x^{T}Qx + \mu^{T}R\mu$ is the utility function and $U(0, 0) = 0$; $Q \in \mathbb{R}^{12 \times 12}$ and $R \in \mathbb{R}^{m \times m}$ are positive definite matrices; $\tilde{f}$ is the approximate actuators failures $f$; $\gamma$ is a discount factor and $0 \leq \gamma < 1$; $\beta$ is a positive constant.

**Definition 1** A control law $\mu$ is defined as an admissible control policy for the error tracking system (10) with $f = 0$, if $\mu$ is continuous on a set $\Omega \subset \mathbb{R}^{12}$ and can stabilize the error tracking system (10) with $f = 0$, $\mu(0) = 0$ and $V_{1}(x_{0}, 0)$ is finite for all $x_{0} \in \Omega$.

Based on the optimal control theory, the performance index function (11) is a Lyapunov function and satisfies as follows:

$$
0 = \beta \tilde{f}^{T} \tilde{f} + U(x, \mu) + (\nabla V_{1}(x, \mu))^{T}(\varpi(x) + \rho(x)\mu) - \gamma V_{1}(x, \mu)
$$

(12)

where $V_{1}(0, 0) = 0$ and $\nabla V_{1}(x, \mu)$ is the partial derivative of $V_{1}(x, \mu)$ with respect to $x$, $\nabla V_{1}(x, \mu) = \frac{\partial V_{1}(x, \mu)}{\partial x}$.

Then, the Hamiltonian function is defined as follows:

$$
H(x, \mu, \nabla V_{1}(x, \mu)) = \beta \tilde{f}^{T} \tilde{f} + U(x, \mu) + (\nabla V_{1}(x, \mu))^{T}(\varpi(x) + \rho(x)\mu)
$$

(13)

$$
-\gamma V_{1}(x, \mu)
$$

The optimal cost function is defined as follows:

$$
V_{1}^{*}(x, \mu) = \min_{\mu \in \Psi(\Omega)} \int_{t}^{\infty} e^{\gamma(t-\sigma)}(\beta \tilde{f}^{T} \tilde{f} + U(x(\sigma), \mu(\sigma)))d\sigma
$$

(14)

where $\delta_{2}$ is a positive constant.
The optimal cost function (14) satisfies the HJB equation, then

$$0 = \min_{\mu} H(x, \mu, \nabla V^*_1(x, \mu)) \quad (15)$$

The optimal control is expressed as follows:

$$\mu^*(x) = -\frac{1}{2} R^{-1} \rho^T(x) \nabla V^*_1(x, \mu) \quad (16)$$

The PI scheme is designed as shown in Algorithm 1.

**Algorithm 1** Online PI

**Step 1:** Select an initial admissible control policy $\mu^{(0)}$ and a positive constant $\epsilon$ and an initial performance index function $\nabla V^0_1(x, \mu^{(0)}) = 0$;

**Step 2:** Solve $V^i_1$ according to

$$0 = \beta \hat{f}^T \hat{f} + U(x, \mu^i) + (\nabla V^i_1(x, \mu^{(i)}))^T (w(x) + \rho(x)\mu^{(i)}) - \gamma V^{i-1}_1(x, \mu^{(i-1)}) ;$$

**Step 3:** Update the control policy with

$$\mu^{(i+1)} = -\frac{1}{2} R^{-1} \rho^T(x) \nabla V^1_1(x, \mu^i) ;$$

**Step 4:** if $\|V^{i+1}_1(x, \mu^{(i+1)}) - V^{(i)}_1(x, \mu^{(i)})\| \leq \epsilon$, stop the iterations; else return to Step 2.

3 Fault-tolerant ADP tracking controller design via neural network observer

3.1 Problem transformation

The structural diagram of neural network observer based fault-tolerant ADP control scheme is shown in Fig. 2.
Assumption 2 The approximate error of actuators faults $e_f = f - \hat{f}$ satisfies that $\|e_f\| \leq \delta_3$, where $\delta_3$ is a positive constant.

Lemma 1 (Zhao et al. (2017, 2016)) With Assumption 1, 2 and the control policy (16) for error tracking system (10) with $f = 0$, the continuously differentiable function $V_i(x, \mu)$ is a Lyapunov function if the conditions $\beta \geq \gamma \delta_1^{-2} \delta_2 \lambda_{\text{max}}(R^{-1}) + \lambda_{\text{max}}(R)$ and $\|x\| \geq \sqrt{(\gamma \delta_1^{-2} \delta_2 \lambda_{\text{max}}(R^{-1}) + \lambda_{\text{max}}(R)) (2\delta_1 + \delta_3)}$ hold. So, the optimal control law (16) is a solution to the error tracking system (10) with $f \neq 0$ and error tracking system (10) with $f \neq 0$ is UUB.

Proof The derivative of $V_i(x, \mu^*)$ is given as follows:

$$\dot{V}_i(x, \mu^*) = (\nabla V_i(x, \mu^*)^T \dot{x}$$

From equation (15) we have

$$(\nabla V_i(x, \mu^*)^T (\nabla x + \rho(x)\mu^*) = -\beta \hat{f}^T \hat{f} - U(x, \mu^*) + \gamma V_i(x, \mu^*)$$

Substituting equation (18) into equation (17), we can get

$$\dot{V}_i(x, \mu^*) = -\beta \hat{f}^T \hat{f} - U(x, \mu^*) + \gamma V_i(x, \mu^*) - (\nabla V_i(x, \mu^*)^T \rho(x)f$$

Hence, if the following conditions hold,

$$\begin{cases} 
\beta \geq \gamma \delta_1^{-2} \delta_2 \lambda_{\text{max}}(R^{-1}) + \lambda_{\text{max}}(R) \\
\|x\| \geq \sqrt{(\gamma \delta_1^{-2} \delta_2 \lambda_{\text{max}}(R^{-1}) + \lambda_{\text{max}}(R)) (2\delta_1 + \delta_3)} 
\end{cases}$$

$\dot{V}_i(x, \mu) \leq 0$ and $V_i(x, \mu)$ is a Lyapunov function.

The error tracking system (10) is UUB. This completes the proof.
3.2 Design of neural-network observer

For the error tracking system (10), we developed a radial basis function (RBF) neural network to approximate the actuators faults.

\[ f = -(W_0 \varphi(x) + \varepsilon_0) \]  

(21)

where \( W_0^T \in \mathbb{R}^{l_0} \) is the ideal weight; \( \varphi_0(x) \in \mathbb{R}^{l_0} \) is the activation function; \( l_0 \) is the neurons number of the hidden layer; \( \varepsilon_0 \) is the approximation error.

Substituting equation (21) into error tracking system (10), we can get

\[ \dot{x} = \varpi(x) + \rho(x)(\mu + W_0 \varphi_0(x) + \varepsilon_0) \]  

(22)

Then the neural-network faults observer is designed as follows:

\[ \dot{\hat{x}} = \varpi(\hat{x}) + \rho(\hat{x})(\mu + \hat{W}_0 \varphi_0(\hat{x})) + L(x - \hat{x}) \]  

(23)

where \( \hat{x} \) the approximation of \( x \); \( \hat{W}_0 \) is the approximation of \( W_0 \); \( L \in \mathbb{R}^{12 \times 12} \) is the positive matrix.

The weight vector \( \hat{W}_0 \) should be updated as

\[ \dot{\hat{W}}_0 = -\varrho_0 \rho_T(\hat{x}) e_x \varphi_0^T(\hat{x}) \]  

(24)

where \( e_x = \hat{x} - x \) is the approximation error of \( x \), \( \varrho_0 \) is the learning rate and \( \varrho_0 > 0 \).

Combining equation (22) with equation (23), we can get

\[ \dot{e}_x = \dot{\hat{x}} - \dot{x} \]

\[ = \varpi(\hat{x}) + \rho(\hat{x})(\mu + \hat{W}_0 \varphi_0(\hat{x})) - \varpi(x) - \rho(x)(\mu + W_0 \varphi_0(x) + \varepsilon_0) \]

\[ = -Le_x + (\varpi(\hat{x}) - \varpi(x)) + (\rho(\hat{x}) - \rho(x))\mu + \rho(\hat{x})\hat{W}_0 \varphi_0(\hat{x}) - \rho(x)W_0 \varphi_0(x) - \rho(\hat{x})\hat{W}_0 \varphi_0(\hat{x}) \]

\[ = -Le_x + \varpi + \rho(\hat{x})\hat{W}_0 \varphi_0(\hat{x}) - \rho(x)W_0 \varphi_0(x) - \rho(\hat{x})\hat{W}_0 \varphi_0(\hat{x}) \]

\[ + (\rho(\hat{x})\hat{W}_0 \varphi_0(\hat{x}) - \rho(\hat{x})\hat{W}_0 \varphi_0(x) - \rho(\hat{x})\hat{W}_0 \varphi_0(\hat{x}) - \rho(\hat{x})\hat{W}_0 \varphi_0(x) - \rho(\hat{x})\hat{W}_0 \varphi_0(\hat{x}) \]

\[ = -Le_x + \varpi + \rho(\hat{x})\hat{W}_0 \varphi_0(\hat{x}) + \rho(x)\hat{W}_0 \varphi_0(x) + \rho(\hat{x})\hat{W}_0 \varphi_0(\hat{x}) \]

\[ = -Le_x + \varpi + \rho(\hat{x})\hat{W}_0 \varphi_0(\hat{x}) + \rho(x)\hat{W}_0 \varphi_0(x) + \rho(\hat{x})\hat{W}_0 \varphi_0(\hat{x}) + \rho(x)\hat{W}_0 \varphi_0(x) - \rho(\hat{x})\hat{W}_0 \varphi_0(\hat{x}) \]

(25)

where \( \varpi = \varpi(\hat{x}) - \varpi(x); \hat{\rho} = \rho(\hat{x}) - \rho(x); \hat{\varphi}_0 = \varphi_0(\hat{x}) - \varphi_0(x); \hat{W}_0 = \hat{W}_0 - W_0 \)

and \( \hat{W}_0 = \hat{W}_0 \).

**Assumption 3** \( \varpi, \hat{\rho}, \hat{\varphi}_0, \hat{W}_0 \varphi_0(x) \), \( \rho(x)\hat{W}_0 \varphi_0(x) \) and \( \rho(x)\varepsilon_0 \) are norm-bounded as \( \|\varpi\| \leq \delta_4 \), \( \|\hat{\rho}\| \leq \delta_5 \), \( \|\hat{\varphi}_0\| \leq \delta_6 \), \( \|\hat{W}_0 \varphi_0(x)\| \leq \delta_7 \) and \( \|\rho(x)\varepsilon_0\| \leq \delta_8 \); \( \delta_4, \delta_5, \delta_6, \delta_7 \) and \( \delta_8 \) are positive constants.

**Theorem 1** With Assumptions 1, 3, the updating law (24) of weight vector (23) can guarantee \( e_x \) to be UUB based on the neural network observer.
Proof Select an Lyapunov function as
\[ V_2 = \frac{1}{2} T_x e_x + \frac{1}{2g_0} tr[W_0^T \dot{W}_0] \] (26)

Substituting equation (25) into the time derivative of equation (26), we can get
\[
\dot{V}_2 = e_x^T \dot{e}_x + \frac{1}{g_0} tr[\dot{W}_0^T \dot{W}_0]
\]
\[
= e_x^T (-L e_x + \tilde{\omega} + \tilde{\rho} \mu + \rho(\dot{x}) \tilde{W}_0 \tilde{v}_0(\dot{x}) + \tilde{\rho} W_0 \tilde{v}_0(\dot{x}) + \rho(x) W_0 \tilde{v}_0 - \rho(x) e_0)
\]
\[
+ tr[(\mu^T(\dot{x}) e_x \phi_0(\dot{x}))^T \dot{W}_0]
\]
\[
= e_x^T (-L e_x + \tilde{\omega} + \tilde{\rho} \mu + \rho(\dot{x}) \tilde{W}_0 \tilde{v}_0(\dot{x}) + \tilde{\rho} W_0 \tilde{v}_0(\dot{x}) + \rho(x) W_0 \tilde{v}_0 - \rho(x) e_0)
\]
\[
- tr[\phi_0(\dot{x}) e_x^T \rho(\dot{x}) W_0]
\]
\[
= e_x^T (-L e_x + \tilde{\omega} + \tilde{\rho} \mu + \rho(\dot{x}) \tilde{W}_0 \tilde{v}_0(\dot{x}) + \tilde{\rho} W_0 \tilde{v}_0(\dot{x}) + \rho(x) W_0 \tilde{v}_0 - \rho(x) e_0)
\]
\[
\leq -\lambda_{\min}(L) e_x^T e_x + e_x^T (\tilde{\omega} + \tilde{\rho} \mu + \rho W_0 \tilde{v}_0(\dot{x}) + \rho(x) W_0 \tilde{v}_0 - \rho(x) e_0)
\]
\[
\leq -\lambda_{\min}(L) e_x^T e_x + \frac{4}{\lambda_{\min}} (\|\tilde{\omega}\|^2 + \|\tilde{\rho} \mu\|^2 + \|\tilde{\rho} W_0 \tilde{v}_0(\dot{x})\|^2 + \|\rho(x) W_0 \tilde{v}_0 - \rho(x) e_0\|^2)
\]
\[
\leq -\lambda_{\min}(L) e_x^T e_x + \frac{4}{\lambda_{\min}} (\delta_1^2 + \delta_2^2 + \delta_3^2 + \delta_4^2)
\] (27)

We can conclude that \(\dot{V}_2 < 0\) if \(e_x\) satisfies \(\|e_x\| > \frac{\lambda_{\min}}{4} \sqrt{\delta_1^2 + \delta_2^2 + \delta_3^2 + \delta_4^2}\).
Based on the Lyapunov stability theorem, \(e_x\) is guaranteed to be UUB. This completes the proof.

3.3 Design of critic neural network

The ADP controller consists of critic neural network and action neural network. The critic neural network is utilized to approximate \(V^*_1(x, \mu^*).\)
\[
V_3(x, \mu) = W_c c_\varphi(x, \mu) + \varepsilon_c
\] (28)
where \(W_c \in \mathbb{R}^{l1}\) is the ideal weight; \(c_\varphi(x, \mu) \in \mathbb{R}^{l1}\) is the activation function;
\(l_1\) is the neurons number of the hidden layer; \(\varepsilon_c\) is the approximation error.

The derivative of the cost function \(V_3(x, \mu)\) is given as follows
\[
\nabla V_3(x, \mu) = (\nabla \varphi_\mu x, \mu)) W_c^T + \nabla \varepsilon_c
\] (29)
where \(\nabla \varphi_\mu x, \mu = \frac{\partial \varphi(x, \mu)}{\partial x}\) and \(\nabla \varepsilon_c = \frac{\partial \varepsilon_c}{\partial x}\).
Substituting equation (29) into equation (12), we can obtain
\[
0 = \beta \tilde{f}^T \tilde{f} + U(x, \mu) + (W_c \nabla \varphi(x, \mu) + \nabla \varepsilon_c)((\varphi(x) + \rho(x)(\mu - \hat{f})\n- \gamma (W_c \varphi(x, \mu) + \varepsilon_c)
\] (30)
Then the Hamiltonian function can be expressed as follows:

\[
H(x, \mu, W_c) = \beta \hat{f}^T \hat{f} + U(x, \mu) + W_c \nabla \varphi_c(x, \mu)(\varpi(x) + \rho(x)(\mu - \hat{f}))
- \gamma W_c \varphi_c(x, \mu)
= e_c
\]

(31)

where \( e_c \) is the residual error.

Then, \( V_3(x, \mu) \) is approximated as follows:

\[
\hat{V}_3(x, \mu) = \hat{W}_c \varphi_c(x, \mu)
\]

(32)

where \( \hat{W}_c \) is the approximation of \( W_c \).

The derivative of \( \hat{V}_3(x) \) can be expressed as follows:

\[
\nabla \hat{V}_3(x, \mu) = (\nabla \varphi_c(x, \mu)) \hat{W}_c^T
\]

(33)

Then, the approximate Hamiltonian function can be expressed as follows:

\[
\hat{H}(x, \mu, \hat{W}_c) = \beta \hat{f}^T \hat{f} + U(x, \mu) + \hat{W}_c \nabla \varphi_c(x, \mu)(\varpi(x) + \rho(x)(\mu - \hat{f}))
- \gamma \hat{W}_c \varphi_c(x, \mu)
= \hat{e}_c
\]

(34)

Given any admissible control policy \( \mu \), it is desired to select \( \hat{W}_c \) to minimize the squared residual error \( E_c(\hat{W}_c) \) as

\[
E_c(\hat{W}_c) = \frac{1}{2} \hat{e}_c^T \hat{e}_c
\]

(35)

The weight update law for the critic neural network is given as follows

\[
\dot{\hat{W}}_c = -\frac{\varrho_1 \hat{e}_c \varsigma_1^T}{(1 + \varsigma_1^T \varsigma_1)^2}
\]

(36)

where \( \varrho_1 \) is the learning rate of critic neural network and \( \varrho_1 \) satisfies that \( \varrho_1 > 0; \varsigma_1 = \nabla \varphi_c(x, \mu)(\varpi(x) + \rho(x)(\mu - \hat{f})) - \gamma \varphi_c(x, \mu) \) and \( \varsigma_1 \in \mathbb{R}^{l_1} \).

The approximate weight error of critic neural network is defined as \( \tilde{W}_c = \hat{W}_c - W_c \). Then, equation (36) can be transformed as follows:

\[
\dot{\tilde{W}}_c = -\frac{\varrho_1}{(1 + \varsigma_1^T \varsigma_1)^2}(\tilde{W}_c \varsigma_1 - \nabla \varepsilon_c(\varpi(x) + \rho(x)(\mu - \hat{f})) - \gamma \varepsilon_c)\varsigma_1^T
\]

(37)

**Assumption 4** \( ||\nabla \varepsilon_c(\varpi(x) + \rho(x)(\mu - \hat{f})) + \gamma \varepsilon_c|| \leq \delta_0 \) and \( \varsigma_{\min} \leq ||\varsigma_1|| \leq \varsigma_{\max} \), where \( \delta_0 \), \( \varsigma_{\min} \) and \( \varsigma_{\max} \) are positive constants.

**Theorem 2** The approximate weight error is UUB, if the weight of the critic neural network is updated by (37).
Proof Select an Lyapunov function as

\[ V_4 = \frac{(1 + \varsigma l^T \varsigma_1)^2}{2\dot{\rho}_1} \mathbf{W}_c \mathbf{W}_c^T \]  

Then, the time derivative of \( V_4 \) is

\[ \dot{V}_4 = \frac{(1 + \varsigma l^T \varsigma_1)^2}{2\dot{\rho}_1} \dot{\mathbf{W}}_c \mathbf{W}_c^T \]
\[ = -(\mathbf{W}_c \varsigma_1 - \nabla \varepsilon_c(\varphi(x) + \rho(x)(\mu - \hat{f})) - \gamma \varepsilon_c \varsigma_1^T \mathbf{W}_c^T \]
\[ = -\mathbf{W}_c \varsigma_1^T \mathbf{W}_c^T + (\nabla \varepsilon_c(\varphi(x) + \rho(x)(\mu - \hat{f})) + \gamma \varepsilon_c \varsigma_1^T \mathbf{W}_c^T \]
\[ \leq -\frac{1}{2} \| \mathbf{W}_c \varsigma_1 \|^2 + \frac{1}{2} \| \nabla \varepsilon_c(\varphi(x) + \rho(x)(\mu - \hat{f})) + \gamma \varepsilon_c \|_2^2 \]
\[ \leq -\frac{1}{2} \| \mathbf{W}_c \varsigma_1 \|^2 + \frac{1}{2} \| \delta \|_2^2 \]

Hence, \( \dot{V}_4 < 0 \) if \( \| \mathbf{W}_c \| > \| \frac{\varsigma_1}{\varepsilon} \| \). The approximate weight error is UUB, according to the Lyapunov stability theorem. This completes the proof.

3.4 Design of action neural network

The optimal control \( \mu^* \) is approximated by the action neural network as

\[ \mu = \mathbf{W}_a \varphi_a(x) + \varepsilon_a \]  

where \( \mathbf{W}_a \in \mathbb{R}^{l_2} \) is the ideal weight; \( \varphi_a(x) \in \mathbb{R}^{l_2} \) is the activation function; \( l_2 \) is the neurons number of the hidden layer; \( \varepsilon_a \) is the approximation error.

Because the ideal weight \( \mathbf{W}_a \) is unknown, \( \mu^* \) is approximated as follows:

\[ \hat{\mu} = \hat{\mathbf{W}}_a \varphi_a(x) \]  

where \( \hat{\mathbf{W}}_a \) is the estimate of \( \mathbf{W}_a \).

The approximate feedback error used for training action neural network is defined as the difference between the feedback control input applied to the error tracking system (10) and the optimal control \( \mu^* \) as

\[ \hat{\varepsilon}_a = \hat{\mathbf{W}}_a \varphi_a(x) + \frac{1}{2} R^{-1} \rho^T(x)(\nabla \varphi_a(x, \mu))^T \hat{\mathbf{W}}_c^T \]  

The action neural network is defined to minimize the objective function as

\[ E_a(\hat{\mathbf{W}}_a) = \frac{1}{2} \hat{\varepsilon}_a^T \hat{\varepsilon}_a \]  

The weight updating law for the action neural network is given as follows

\[ \dot{\hat{\mathbf{W}}}_a = -\varrho_2 \hat{\varepsilon}_a \varphi_a^T(x) \]  

where \( \varrho_2 \) is the learning rate of action neural network and \( \varrho_2 > 0 \).

According to equations (16), (29) and (40), we have

\[ 0 = \mathbf{W}_a \varphi_a(x) + \varepsilon_a + \frac{1}{2} R^{-1} \rho^T(x)(\nabla \varphi_a(x, \mu))^T \mathbf{W}_c^T + \nabla \varepsilon_c \]
The approximate weight error of action neural network is defined as $\hat{W}_a = W_a - \hat{W}_a$. Then, equation (44) can be transformed as follows:

$$\dot{\hat{W}}_a = -\eta_{\hat{a}}(\hat{W}_a \psi_a(x) + \frac{1}{2} R^{-1} \rho^T(x)(\nabla \varphi_c(x, \mu))^T \hat{W}_c - \varepsilon_a - \frac{1}{2} R^{-1} \rho^T(x)(\nabla \varepsilon_c) \varphi_a^T(x))$$  \hfill (46)

**Assumption 5** $\| R^{-1} \rho^T(x)(\nabla \varphi_c(x, \mu)) \| \leq \delta_{10}$, $\| \varepsilon_a + \frac{1}{2} R^{-1} \rho^T(x)(\nabla \varepsilon_c) \| \leq \delta_{11}$, and $\varphi_{a, \min} \leq \| \varphi_a(x) \| \leq \varphi_{a, \max}$, where $\delta_{10}, \delta_{11}, \varphi_{a, \min}$ and $\varphi_{a, \max}$ are positive constants.

**Theorem 3** The approximate weight error is UUB, if the weight of the action neural network is updated by (46).

**Proof** Select an Lyapunov function as

$$V_5 = \frac{1}{2\eta^2} \hat{W}_a \hat{W}_a^T$$  \hfill (47)

Then, the time derivative of $V_5$ is

$$\dot{V}_5 = \frac{1}{2\eta^2} \dot{\hat{W}}_a \hat{W}_a^T + \frac{1}{2\eta} \hat{W}_a \dot{\hat{W}}_a^T$$

$$\dot{V}_5 = -\frac{1}{2} \hat{W}_a \dot{\hat{W}}_a + \frac{1}{2} R^{-1} \rho^T(x)(\nabla \varphi_c(x, \mu))^T \hat{W}_c^T$$

$$\leq -\frac{1}{2} \| W_a \| \| \dot{W}_a \| + \frac{1}{2} \| R^{-1} \rho^T(x)(\nabla \varphi_c(x, \mu))^T \| \hat{W}_c^T \| ^2$$

$$\leq -\frac{1}{2} \| W_a \| \| \dot{W}_a \| + \frac{1}{2} \| R^{-1} \rho^T(x)(\nabla \varphi_c(x, \mu))^T \| \| \hat{W}_c \| ^2$$

Hence, $\dot{V}_5 < 0$ if $\| \hat{W}_a \| > \sqrt{\frac{\delta_{11}^2 \| W_a \| ^2 + 2\delta_{11}^2}{\varphi_{a, \min}}}$.

3.5 Stability analysis

**Assumption 6** $\| \varphi(x) \| \leq \delta_{12}$ and $\| \rho(x) \| \leq \delta_{13}$, where $\delta_{12}$ and $\delta_{13}$ are positive constants.

**Theorem 4** With the performance index function (11), the error tracking system (10) can be guaranteed to be UUB by the approximate fault-tolerant tracking control policy (41).

**Proof** Select an Lyapunov function as

$$V_6 = \frac{1}{2} \dot{x}^T x + V_1^*$$  \hfill (49)
Then, the time derivative of $V_6$ is

$$
V_6 = x^T \dot{x} + (\nabla V_1^T)^T \dot{x} \\
= x^T (\varpi(x) + \rho(x)(\mu - f)) + (\nabla V_1^T(\varpi(x) + \rho(x)(\mu - f)) \\
= x^T \varpi(x) + x^T \rho(x) \mu - x^T \rho(x)f - (\nabla V_1^T) \rho(x)f \\
+ (\nabla V_1^T)^T(\varpi(x) + \rho(x)\mu)
$$

(50)

According to equations (13), (15), the equation (50) can be transformed as follows:

$$
\dot{V}_6 = x^T \varpi(x) + x^T \rho(x) \mu - x^T \rho(x)f - (\nabla V_1^T) \rho(x)f - x^T Qx - \mu^T R \mu \\
- \beta \dot{f} f + \gamma V_4^* \\
= x^T \varpi(x) + x^T \rho(x) \mu - x^T \rho(x)f + 2 \mu^T R f - x^T Qx - \mu^T R \mu \\
- \beta \dot{f} f + \gamma V_4^* \\
\leq \frac{3}{2} x^T x + \frac{1}{2} \varpi^T(x) \varpi(x) + \frac{1}{2} \mu^T \rho^T(x) \rho(x) \mu + \frac{1}{2} f^T \rho^T(x) \rho(x) f \\
+ \mu^T \mu + f^T R^T R f - x^T Qx - \mu^T R \mu - \beta \dot{f} f + \gamma \delta_2 \\
\leq -\widehat{\lambda}_{\min}(Q) \frac{3}{2} ||x||^2 - (\widehat{\lambda}_{\min}(Q) - \frac{1}{2} \delta_{13}^2 - 1)||\mu||^2 - (\beta - \frac{1}{2} \delta_{13}^2 - \beta \dot{f} f + \gamma \delta_2) \\
- \lambda_{\max}(R') ||f||^2 + (\frac{1}{2} \delta_{13}^2 + \lambda_{\max}(R'))(\delta_1 + \delta_3) \delta_3 + \frac{1}{2} \delta_{12}^2 + \gamma \delta_2
$$

(51)

where $R' = R^T R$.

Hence, $\dot{V}_6 < 0$ if $(\lambda_{\min}(Q) - \frac{3}{2}) > 0$, $(\lambda_{\min}(Q) - \frac{1}{2} \delta_{13}^2 - 1) \geq 0$, $(\beta - \frac{1}{2} \delta_{13}^2 - \beta \dot{f} f + \gamma \delta_2) > 0$ and $||x|| > \sqrt{\frac{(\frac{1}{2} \delta_{13}^2 + \lambda_{\max}(R'))(\delta_1 + \delta_3) \delta_3 + \frac{1}{2} \delta_{12}^2 + \gamma \delta_2}{\lambda_{\min}(Q) - \frac{3}{2}}}$, The error tracking system (10) is UUB, according to the Lyapunov stability theorem. This completes the proof.

4 Simulation results

In order to show the effectiveness of the proposed fault-tolerant tracking control based on ADP, two simulation examples are given in this section. According to the kinematic and dynamic model of underactuated AUV (1) with the conditions that are $\eta(4) = 0$ and $\xi(4) = 0$, the matrices $M$, $C(\xi)$, $D(\xi)$ and $g(\eta)$, $J$ are described as follows:

$$
M = \begin{bmatrix}
215 & 0 & 0 & 0 & 0 & 0 \\
0 & 265 & 0 & 0 & 0 & 0 \\
0 & 0 & 265 & 0 & 0 & 0 \\
0 & 0 & 0 & 80 & 0 & 0 \\
0 & 0 & 0 & 0 & 80 & 0 \\
0 & 0 & 0 & 0 & 0 & 265
\end{bmatrix}
$$

(52)

$$
C(\xi) = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & -265w \\
0 & 0 & 0 & 265w & -265v & 0 \\
0 & 0 & 0 & -215u & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
-50w & 0 & 0 & 0 & 0 & 0 \\
50v & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
$$

(53)
\[ D(\xi) = \begin{bmatrix} D_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & D_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & D_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & D_4 & 0 & 0 \\ 0 & 0 & 0 & 0 & D_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & D_6 \end{bmatrix} \]  
\tag{54}

where \( D_1 = 70 + 100|u|; \) \( D_2 = 100 + 200|v|; \) \( D_3 = 100 + 200|w|; \) \( D_4 = 0; \) \( D_5 = 50 + 100|q|; \) and \( D_6 = 50 + 100|r|. \)

\[ g(\eta) = [0 \ 0 \ -(1822.25 - G)\cos(\theta) \ 0 \ (18.2225 - 0.01G)\sin(\theta) \ 0]^T \]  
\tag{55}

where \( G \) is the gravity.

\[ J = \begin{bmatrix} c\theta c\psi - s\psi s\theta & c\theta s\psi s\psi & 0 & 0 \end{bmatrix} \]  
\tag{56}

where \( c\theta = \cos\theta; \) \( s\theta = \sin\theta; \) \( s\psi = \sin\psi; \) \( c\psi = \cos\psi \)

### 4.1 Example one without actuators faults

Given \( \tau_f = 0, \ \varphi_0 = 0.1, \ \varphi_1 = 0.02, \ \varphi_2 = 0.04, \ \gamma = 0.3, \ B = 1822.25, \ \beta = 0.15 \) and \( \tau_d = [500 \ \ 0 \ \ 0 \ \ 200 \ \ 10]^T, \) the simulation results are given as follows.
Fig. 3 and Fig.4 show the tracking error of desired position and the tracking error of desired velocity with respect to inertial coordinate system compared with the existing method (Zhao et al (2017, 2016)) respectively. The tracking trajectory is shown in Fig.5. The method proposed in this work has received almost the same results with the existing method (Zhao et al (2017, 2016)). From Fig.3 and Fig.4, we can know that the error tacking system (10) is bounded stable. The absolute value of the tracking error of desired position is no more than the threshold value 0.2. The absolute value of the tracking error of desired velocity is no more than the threshold value 0.05. From Fig.5, we know that the value of the error trajectory between the desired trajectory and the simulation trajectory with the method proposed in this work is no bigger than 0.1m.

4.2 Example two with actuators faults

In this simulation example, we used the parameters values of example one except for $\tau_f = 0.1(\tau_c + \tau_d)$ and $\tau_f = 0.2(\tau_c + \tau_d)$ respectively. The simulation results compared with the existing method (Zhao et al (2017, 2016)) are given as follows.
Fig. 5 AUV trajectory

(a) Tracking error of desired position with $\tau_f = 0.1(\tau_e + \tau_d)$
(b) Tracking error of desired position with $\tau_f = 0.2(\tau_e + \tau_d)$

Fig. 6 Tracking error of desired position

Fig.6 and Fig.7 show the tracking error of desired position and the tracking error of desired velocity with respect to inertial coordinate system respectively compared with the existing method (Zhao et al (2017, 2016)). From Fig.7, we know that the jitter happens with the existing method (Zhao et al (2017, 2016)) from 0 second to 15 second, and when actuators faults $\tau_f$ is $0.2(\tau_e + \tau_d)$, the jitter is bigger. The method proposed in this paper has received better results without jitter. The tracking error of desired position and the tracking error of desired velocity are tracking system are bounded.

Fig.8 and Fig.9 give the estimated actuators faults based on the RBF neural network when the values of $\tau_f$ are $0.1(\tau_e + \tau_d)$ and $0.2(\tau_e + \tau_d)$ respectively. With the actuators faults, the jitter happened in the estimation of actuators.
faults with existing method (Zhao et al. (2017, 2016)). When the actuators faults became bigger, the jitter became bigger.

Fig. 10 shows the tracking trajectories with $\tau_f = 0.1(\tau_e + \tau_d)$ and $\tau_f = 0.2(\tau_e + \tau_d)$. From the simulation results, we know that the value of the error trajectory between the desired trajectory and the simulation trajectory with
5 Conclusion

ADP tracking technique via neural network observer based on the error tracking system with actuators faults has been designed. The simulation examples are developed without actuators faults and with actuators faults respectively for the development of the fault-tolerant tracking control scheme to achieve better tracking performance. When the actuators faults happened, the jitter appeared with the existing method (Zhao et al (2017, 2016)). The jitter will become bigger when the actuators faults become bigger and make the system...
unstable. However, the method proposed in this work make the error tracking
system bounded stable with the actuators faults based on the Lyapunov sta-
bility. Simulation results have shown excellent performance of the closed-loop
system compared with the existing method (Zhao et al (2017, 2016)).

References

control for auv via adp with rudders faults and ocean current disturbance.
Neurocomputing 411:442–454
based on particle swarm optimization algorithm for path planning of au-
tonomous underwater vehicle. J Amb Intel Hum Comp 11:3349–3354
Che G, Liu L, Yu Z (2019b) Nonlinear trajectory-tracking control for au-
tonomous underwater vehicle based on iterative adaptive dynamic program-
Gong L, Wang Q, Dong C (2019) Spacecraft output feedback attitude control
based on extended state observer and adaptive dynamic programming. J
Frankl Inst 356:4971–5000
mode output feedback. Appl Math Comput 359:433–455
control study of an underwater robot with four rotatable thrusters. Ocean
Eng 197, DOI 10.1016/j.oceaneng.2020.106929
unknown nonlinear discrete-time systems. IEEE T Neural Netw Learn Syst
26(1):140–151
Li J, Du J, Sun Y, Lewis F (2019) Robust adaptive trajectory tracking control
of underactuated autonomous underwater vehicles with prescribed perfor-
ance. Int J Robust Nonlin 29(14):4629–4643
fault reconstruction for autonomous underwater vehicle. Ocean Eng 155:10–
23
underwater vehicles based on fuzzy re-planning of a local desired trajectory.
IEEE T Veh Technol 68(12):11657–11667
Liu Y, Zhang H, Yu R, Qu Q (2018b) Data-driven optimal tracking control
for discrete-time system with delays using adaptive dynamic program-
graw text content


