Academic Journal of Science and Technology
ISSN: 2771-3032 | Vol. 14, No. 3, 2025

Adaptive Control Strategy for Oscillating Water Column Wave Energy Device Based on Deep Reinforcement Learning

Yuxuan Liu

Dalian University of Technology, Dalian, China

Abstract: This paper proposes an adaptive control strategy for oscillating water column (OWC) wave energy devices based on deep reinforcement learning (DRL). By combining the feature extraction capability of deep neural networks with the decision optimization ability of reinforcement learning, a system is designed to dynamically adjust the turbine control strategy under varying wave conditions. Experimental results show that the proposed method effectively improves the energy conversion efficiency of the device, enhances its stability under extreme sea conditions, and significantly optimizes the system's long-term operational performance. The study demonstrates that deep reinforcement learning provides a feasible solution for intelligent control of wave energy devices.

Keywords: Deep Reinforcement Learning; Oscillating Water Column; Wave Energy; Adaptive Control; Energy Conversion Efficiency.

1. Modeling and Analysis of Oscillating Water Column Wave Energy Device

(A) Basic Structure and Working Principle of Oscillating Water Column Wave Energy Device

1. Structure and Mechanism of the Oscillating Water Column

An Oscillating Water Column (OWC) wave energy device is a system that utilizes the up-and-down motion of water columns induced by ocean waves to drive a gas turbine for power generation. The basic structure typically includes a closed water tank, with a turbine connected to the outside air via a gas pipeline. The upper part of the tank is connected to the atmosphere, while the lower part is in contact with seawater. When waves arrive, the seawater causes vertical oscillations, which drive the water level inside the tank to fluctuate. These fluctuations compress and expand the air inside the tank, thereby rotating the gas turbine and converting mechanical energy into electrical energy.

2. Wave Energy Capture and Conversion

The working principle of the OWC device is based on the longitudinal oscillations of the waves. These oscillations cause periodic changes in the water's height, generating pressure fluctuations above the water column. These gas pressure variations drive the turbine, which is then converted into electrical energy through a generator. The OWC device has good adaptability in capturing wave energy, especially in areas with fluctuating sea conditions. The core of the device is to effectively convert the energy of the water column's oscillations into mechanical energy, which is ultimately captured and converted into usable energy[1].

3. Energy Conversion Efficiency and Influencing Factors

Energy conversion efficiency is an important indicator of the performance of the OWC device. Factors influencing energy conversion efficiency include wave height, frequency, wave directionality, and the design parameters of the device itself. For example, when the wave period matches the resonance frequency of the OWC device, a higher energy conversion efficiency can be achieved, while frequency mismatches lead to energy loss. The size of the water tank, the design of the turbine, and the configuration of the air intake and exhaust ports are also key factors affecting energy conversion efficiency. Additionally, the device's adaptability to various sea conditions directly influences its long-term energy output.

(B) Dynamic Model of the Oscillating Water Column System

1. System Dynamic Modeling Method

The dynamic modeling of the OWC wave energy device is usually achieved by combining the fluid dynamics equations governing the movement of water with the gas dynamics equations describing pressure changes and turbine rotation. The water movement equations are based on the fundamental principles of fluid mechanics, considering the propagation and reflection of waves within the device. The gas dynamics equations describe the pressure variations in the enclosed container and the rotational motion of the turbine. A common approach is to combine fluid dynamics and aerodynamics models for accurate simulation.

2. Analysis of the System's Nonlinear Characteristics

The OWC system typically exhibits nonlinear characteristics, especially when wave amplitudes are large. The oscillations of the water and the airflow are interrelated, and nonlinear dynamics make the system's response depend not only on the input wave but also on the internal interactions within the system. Due to the variations in wave frequency, amplitude, and the dynamic properties of the water column system, a nonlinear model is essential for accurately simulating the device's actual operation. Analyzing these nonlinear characteristics helps in understanding the performance fluctuations and control challenges encountered during practical operation.

3. Stability and Control Requirements of the Model

The control requirements for the OWC wave energy device mainly focus on how to effectively control the turbine and handle the instability factors in the wave energy capture process. Based on the dynamic model, stability analysis helps evaluate the system's performance under different operating conditions, particularly considering the potential instability caused by nonlinear effects. To achieve efficient energy conversion, a suitable control system must be designed to ensure stable operation of the device under varying wave conditions. This is crucial to prevent structural damage or performance degradation, especially in extreme sea conditions.

(C) Energy Extraction Mechanism and Optimization Model

1. Establishing the Wave Energy Extraction Model

The wave energy extraction model is designed based on the dynamic characteristics of the device and the efficiency of the gas turbine. In the model, factors such as wave period, wave amplitude, and direction are usually considered. A gas pressure fluctuation model within the water column is established to calculate the power that the turbine can extract. The model needs to account for the compression and expansion effects of air inside the water tank under different wave conditions, the response of the turbine, and energy losses. By detailing the wave energy extraction model, the performance of the device can be accurately predicted, providing a theoretical basis for designing the control strategy[2].

2. Parameter Optimization and Control Objective Setting

In practical applications, the performance of the oscillating water column wave energy device depends not only on the characteristics of the waves but also on the design parameters of the device, such as the size of the water tank, turbine design, and the dimensions of the intake and exhaust ports. Therefore, parameter optimization is an essential method for improving energy conversion efficiency. By optimizing control objectives and designing operation strategies that adapt to different wave conditions, the maximum capture of wave energy can be achieved. Control objectives should consider factors such as energy conversion efficiency, system stability, and long-term reliability.

3. Optimization Strategies for Improving Energy Conversion Efficiency

Optimization strategies for improving energy conversion efficiency include several aspects. First, the shape of the water tank and the layout of the turbine should be adjusted to achieve optimal wave energy capture. Second, adaptive control strategies can be employed to dynamically adjust the operating state of the turbine, optimizing airflow and pressure based on real-time changes in the waves. Intelligent control methods, such as deep reinforcement learning, can continuously adjust control strategies to optimize the performance of the device under various sea conditions. Additionally, given the variability of wave energy, predictive control and fault diagnosis methods can be used to maintain and adjust the system, further enhancing efficiency.

2. Basic Principles and Applications of Deep Reinforcement Learning

(A) Overview of Deep Reinforcement Learning

1. Basic Concepts of Reinforcement Learning

Reinforcement Learning (RL) is a paradigm in machine learning where an agent learns decision-making strategies through interactions with an environment. The agent takes actions in the environment and adjusts its strategy based on the feedback (rewards or punishments) from the environment. The core of reinforcement learning is to find the optimal behavior strategy by maximizing cumulative rewards. Unlike supervised learning, RL does not rely on pre-labeled data but learns the best strategy through exploration and trial-and-error.

In the framework of reinforcement learning, the agent’s goal is to learn an optimal strategy such that performing a particular action in a given state will yield the maximum long-term reward. Through interactions with the environment, the agent continually evaluates the outcomes of its actions, adjusting its future decisions to gradually approach the optimal strategy.[3].

2. Advantages of Combining Deep Learning with Reinforcement Learning

Deep Learning (DL) is a learning method based on deep neural networks that can automatically extract features from data through multi-layer structures. In reinforcement learning, traditional RL methods (such as Q-learning) are limited by the scale of state and action spaces, especially in complex, high-dimensional problems where traditional methods are less efficient. By combining deep learning, Deep Reinforcement Learning (DRL) can handle larger and more complex state and action spaces, significantly enhancing the agent's learning capabilities and performance.

The introduction of deep learning allows reinforcement learning to automatically extract features from raw data without the need for manual feature design. With multi-layer processing through deep neural networks, DRL can execute complex decision tasks in complex environments, such as image processing, natural language understanding, and complex control tasks. Therefore, DRL has significant advantages in many fields, especially in high-dimensional and complex decision-making problems.

3. Core Algorithms of Deep Reinforcement Learning

Deep Reinforcement Learning combines the decision-making framework of reinforcement learning with the feature extraction capabilities of deep learning. The core algorithms mainly include the following:

Deep Q-Network (DQN): DQN is an important breakthrough in reinforcement learning. It uses deep neural networks to approximate the Q-value function, solving the bottleneck of traditional Q-learning in high-dimensional state spaces. DQN enhances the training process by using experience replay and target networks, greatly improving Q-learning’s application in complex problems.

Policy Gradient Methods: Policy gradient methods directly optimize the agent's behavior strategy rather than indirectly optimizing the value function. The policy gradient algorithm calculates the gradient of the policy and updates the policy parameters to improve performance. Common policy gradient methods include the REINFORCE algorithm.

Actor-Critic Methods: Actor-Critic methods combine value function methods and policy methods by simultaneously learning a policy (Actor) and a value function (Critic) to improve learning efficiency and stability. The Critic evaluates the current policy, and the Actor improves the policy based on the Critic’s evaluations.

(B) Introduction to Deep Reinforcement Learning Algorithms

1. Q-learning and Deep Q-Network (DQN)

Q-learning is a value iteration algorithm used to solve the optimal policy problem in a Markov Decision Process (MDP). In Q-learning, the agent maintains a Q-value table, where the Q-value represents the expected reward for taking a particular action in a given state. By updating the Q-value table, the agent can learn the optimal policy.

However, Q-learning faces challenges as the dimensions of the state and action spaces grow, making it difficult to manage the Q-value table, especially for continuous state spaces. To address this issue, the Deep Q-Network (DQN) was developed. DQN uses deep neural networks to approximate the Q-value function, enabling it to handle high-dimensional state spaces. The core innovation of DQN lies in the use of experience replay and target networks to stabilize the training process, allowing DQN to make significant progress in complex control tasks.

2. Policy Gradient Methods and Actor-Critic Methods

Policy gradient methods directly optimize the parameters of the policy. Traditional value function methods infer the optimal policy indirectly by learning Q-values, whereas policy gradient methods directly optimize the policy function, compute gradients, and update parameters. By backpropagating the gradients, the agent can effectively improve its behavior strategy. Policy gradient methods are particularly effective for high-dimensional continuous action spaces.

The Actor-Critic method combines value function methods and policy gradient methods, overcoming the limitations of both. In Actor-Critic, the Actor generates the action policy, while the Critic is used to evaluate the quality of the current policy and provides feedback to the Actor. The Actor adjusts its behavior strategy based on the Critic's evaluations, thereby improving overall performance. The advantage of Actor-Critic methods is that they combine policy optimization and value evaluation, making them suitable for complex and dynamic environments[4].

3. Reward Design and Exploration Mechanisms in Deep Reinforcement Learning

Reward design is a critical part of reinforcement learning, directly influencing the agent's learning process and final performance. A well-designed reward function can effectively guide the agent to learn the desired behavior patterns. The reward function needs to account for the characteristics of the task and the objectives, avoiding over-rewarding or over-punishing certain behaviors. Common reward design methods include sparse rewards and dense rewards, where sparse rewards typically require the agent to try continuously over a long period to obtain rewards, while dense rewards provide feedback for each action taken.

Exploration mechanisms are another key element of reinforcement learning. The agent needs to balance exploration (trying new actions) with exploitation (choosing the best-known action). Common exploration mechanisms include the ε-greedy strategy, where the agent has a certain probability of selecting a random action to explore, and the remaining time it chooses the current optimal action. Other exploration strategies, such as the Softmax method, are also widely used in deep reinforcement learning.

(C) Applications of Deep Reinforcement Learning in Control

1. Application of Deep Reinforcement Learning in Robot Control

Deep reinforcement learning has made significant progress in robot control applications. Traditional robot control methods typically rely on manual tuning and predefined rules, whereas deep reinforcement learning autonomously learns control strategies through interactions with the environment. In robot control, deep reinforcement learning is applied to tasks such as autonomous navigation, object manipulation, and path planning. Through deep reinforcement learning, robots can gradually improve their control strategies based on feedback in complex environments and adapt to dynamically changing task requirements.

2. Application of Deep Reinforcement Learning in Smart Grid Management

Deep reinforcement learning is widely used in the management of smart grids for load forecasting, demand response, and energy scheduling. Smart grids feature highly complex systems and uncertainty, making traditional control methods difficult to apply. By using deep reinforcement learning, the smart grid can automatically adjust its operation strategy based on real-time power demand and supply conditions, thereby achieving efficient energy distribution and utilization. Deep reinforcement learning can also optimize power system scheduling, reducing energy consumption and operational costs.

3. Application of Deep Reinforcement Learning in Wave Energy Device Control

In the control of wave energy devices, deep reinforcement learning is used to optimize the energy capture process and improve the stability of the device. The energy conversion efficiency of an oscillating water column wave energy device is influenced by factors such as wave variation, system dynamic response, and turbine control strategy. By introducing deep reinforcement learning, the agent can automatically learn the optimal turbine control strategy under dynamic wave conditions, adjusting the device's working state in real-time to improve energy conversion efficiency and reduce energy loss. Furthermore, deep reinforcement learning can continually optimize the control strategy through feedback mechanisms during the long-term operation of the wave energy device, further enhancing system stability and adaptability.

3. Adaptive Control Strategy Design Based on Deep Reinforcement Learning

(A) Requirements and Goals of Adaptive Control Strategies

1. Basic Principles of Adaptive Control

Adaptive control is a control strategy that automatically adjusts control parameters based on the dynamic characteristics of the system. Unlike traditional fixed-parameter control methods, adaptive control can dynamically adjust its behavior strategy in real-time according to changes in system output and environmental conditions to maintain optimal system performance under uncertainty and varying conditions. In control theory, adaptive control focuses not only on achieving system stability but also on continuously optimizing control parameters through online learning to adapt to unknown or changing environments.

An adaptive control system typically includes system identification, control law adjustment, and feedback mechanisms. The system identification phase involves real-time collection of system input and output data to estimate the system's parameters or model; the control law adjustment phase then adjusts the control strategy based on the information derived from identification, making system behavior align with the desired goals. In the context of deep reinforcement learning, adaptive control models the system using deep learning networks and optimizes the strategy using reinforcement learning algorithms, allowing the control system to adjust and optimize effectively under different operating conditions.

2. Control Requirements of Oscillating Water Column Wave Energy Devices

An oscillating water column wave energy device is a highly wave-dependent energy conversion system. The control objectives include optimizing energy conversion efficiency under varying wave conditions, maintaining long-term stable operation of the device, and preventing damage under extreme sea conditions. The specific control requirements are as follows:

Real-time adaptation to wave variations: Wave height, frequency, and direction constantly change, and traditional control strategies struggle to handle these complex dynamic changes. Deep reinforcement learning can adjust turbine working parameters in real-time according to the current wave conditions, making the system adaptable to wave fluctuations.

Optimization of energy conversion efficiency: The efficiency of wave energy conversion is influenced by wave frequency and amplitude. The control strategy must adjust the turbine speed and other parameters under different wave conditions to maximize energy capture efficiency.

Ensuring system stability and safety: Wave energy devices must withstand extreme wave conditions, such as storms or large fluctuations in sea waves. During these conditions, precise control strategies are necessary to maintain device stability and avoid excessive oscillations or damage.

3. Advantages of Deep Reinforcement Learning in Adaptive Control

Deep reinforcement learning (DRL) offers significant advantages as an adaptive control method. Firstly, DRL can automatically extract key features from raw state data without relying on precise physical models. Secondly, reinforcement learning adjusts control strategies through continuous interaction with the environment, enabling the system to respond flexibly to environmental changes and optimize control decisions via reward and punishment mechanisms.

In wave energy device control, deep reinforcement learning can allow the agent to learn the optimal control strategy for different sea conditions through interaction with the environment. Furthermore, DRL can continually improve strategies through multiple rounds of training, ensuring that wave energy devices maintain high energy conversion efficiency under varying wave conditions and stable operation under extreme waves.

(B) Control Strategy Design Based on Deep Reinforcement Learning

1. Environmental State Modeling and Feature Extraction

The design of control strategies based on deep reinforcement learning starts with modeling the environment. In the oscillating water column wave energy device, the environment state includes variables such as wave height, frequency, speed, and the device's own state, such as turbine speed and water column oscillation amplitude. By monitoring these state variables, the control system can be provided with real-time environmental data.

Deep neural networks, such as Convolutional Neural Networks (CNN) or Recurrent Neural Networks (RNN), can be used for feature extraction, transforming raw environmental data into meaningful high-dimensional feature representations. The goal of feature extraction is to convert dynamic wave changes into inputs that the system can process, allowing the control strategy to react quickly to wave variations. In complex environments, deep reinforcement learning can automatically extract the key factors influencing energy conversion efficiency without the need for manual feature design.

2. Application of Reinforcement Learning Algorithms in Control Strategies

The core of reinforcement learning algorithms is optimizing the agent's behavior strategy through a reward mechanism. For the control of wave energy devices, commonly used reinforcement learning algorithms include Deep Q Networks (DQN), Policy Gradient Methods, and Actor-Critic Methods.

Deep Q Network (DQN): DQN approximates the Q-value function using deep neural networks to evaluate the expected reward for a given action in a particular state. By continually updating the Q-value function, the agent learns the optimal control strategy. The advantage of DQN lies in its ability to handle high-dimensional state spaces and efficiently use past experiences through experience replay during training.

Policy Gradient Methods: Policy gradient methods directly optimize the agent's control policy by calculating gradients to adjust policy parameters. This method is suitable for continuous action spaces and can effectively optimize continuous control tasks, such as adjusting turbine speed or controlling airflow.

Actor-Critic Methods: The Actor-Critic method combines value function and policy function, where the Critic evaluates the current policy's quality and the Actor adjusts its strategy based on the Critic's feedback. This method demonstrates strong stability and training efficiency, making it suitable for complex control tasks.

3. Control Strategy Adjustment and Optimization

The adjustment and optimization of control strategies are central to the adaptive control system based on deep reinforcement learning. By analyzing feedback from the wave energy device in real-time, the control system can adjust the turbine working parameters based on environmental changes. For example, when wave frequency is high, the system can increase the turbine speed; when the wave frequency is low, it can reduce speed to minimize energy loss. Through deep reinforcement learning, the system can automatically adjust control strategies under changing wave conditions, improving energy conversion efficiency.

Additionally, the optimization process considers not only immediate energy capture but also long-term operational stability. Deep reinforcement learning can incorporate device durability and stability into the optimization goal by designing appropriate reward functions, ensuring the control strategy maximizes energy output while maintaining the device's safety.

(C) Training and Optimization of Control Strategies

1. Training Process of Reinforcement Learning Models

The training of reinforcement learning models typically involves the following steps: environment initialization, action selection, reward feedback, and policy updates. Initially, the agent adopts a random control strategy and interacts with the environment. Each interaction generates a combination of state, action, and reward, allowing the agent to accumulate experience and evaluate the effects of different actions.

During training, the agent continually adjusts its behavior strategy through Q-value functions, policy gradients, or

Actor-Critic updates. To enhance training stability and efficiency, experience replay is often used, where past experiences are stored in a memory pool and randomly sampled for updates, breaking data correlations and avoiding overfitting.

2. Design and Adjustment of Reward Functions

The reward function is crucial in determining the agent's behavior. Designing a reasonable reward function helps guide the agent toward optimizing the control strategy. In the control of wave energy devices, the reward function must account for factors like energy conversion efficiency, system stability, and device longevity. For example, the reward can be set based on the turbine's output power, with penalty terms related to the device's health status to prevent excessive oscillations or structural damage.

Adjusting the reward function is also a critical optimization process. During training, the reward function can be dynamically adjusted based on the agent's performance to ensure that optimization goals are effectively achieved.

3. Evaluation and Adjustment of Control Effectiveness

The effectiveness of the control strategy needs to be evaluated through simulation and experimental verification. In the practical application of wave energy devices, comparing the performance of traditional control methods with deep reinforcement learning methods can assess the advantages of DRL control strategies. Evaluation metrics include energy conversion efficiency, device stability, and turbine operational status.

Through multiple rounds of training and evaluation, the agent can gradually optimize the control strategy under various wave conditions. Ultimately, the adaptive control strategy based on deep reinforcement learning will enable efficient operation of the wave energy device while ensuring its stability and long-term reliability in complex environments.

4. Conclusion

The adaptive control strategy based on deep reinforcement learning significantly improves the energy conversion efficiency and stability of the oscillating water column wave energy device. By adjusting turbine control parameters in real time, this method effectively addresses the dynamic changes in wave conditions and optimizes the device's performance over long-term operation, showing broad application prospects.

References

[1] GONÇALVES R A A C,TEIXEIRA P R F,DIDIER E,et al.Numerical analysis of the influence of air compressibility effects on an oscillating water column wave energy converter chamber [J]. Renewable energy, 2020, 153: 1183-1193. DOI: 10.1016/j.renene.2020.02.080.

[2] CUI Y,LIU Z,ZHANG X X,et al.Self-starting analysis of an OWC axial impulse turbine in constant flows: experimental and numerical studies[J].Applied ocean research,2019,82: 458-469.DOI: 10.1016/j.apor.2018.11.014.

[3] LIU Z,JIN J Y,CUI Y,et al.Numerical analysis of impulse turbine for isolated pilot OWC system [J]. Advances in mechanical engineering, 2013.DOI: 10.1155/ 2013/416109.

[4] LIU Z,XU C L,SHI H D,et al.Wave-flume tests of a model-scaled OWC chamber-turbine system under irregular wave conditions [J]. Applied ocean research,2020,99: 102141.DOI: 10.1016/j.apor.2020.102141.