Maximum Power Point Tracking of Photovoltaic System Based on Reinforcement Learning

Kuan-Yu Chou*, Shu-Ting Yang, Chia-Shiou Yang and Yon-Ping Chen
Institute of Electrical and Control Engineering
National Chiao Tung University
1001 University Road, Hsinchu, Taiwan
eric210092@gmail.com
*:corresponding author

Abstract—Maximum power point tracking technique is often used in photovoltaic (PV) system to extract the maximum power at any environment condition. In this paper, a reinforcement learning based variable step size maximum power point tracking (RL MPPT) method is proposed. Q-learning is used as the algorithm of the proposed methods and is implemented by constructing the Q-table (RL-QT MPPT). A Q-network approach (RL-QN MPPT) is also proposed as a more general representation of the RL MPPT method. Implementing of the algorithm doesn’t require the information of the actual PV module in advance, and the proposed system is able to track the MPP offline. With smaller ripples and faster tracking speed, the experiment results of the RL-QT MPPT method and the RL-QN MPPT method are presented.

Keywords—Maximum power point tracking, photovoltaic PV system, reinforcement learning, Q-learning, Q-network

I. INTRODUCTION

Sustainable energy such as solar energy is often seen as one of the solutions to reduce the pollution caused by thermal power generation. A PV module is able to convert the solar energy into electrical energy without generating the greenhouse gases and coal dust, and it is wildly used since the deployment is relatively easy compared to other sustainable energy such as tidal energy and biogas energy. However, the low efficiency is the main drawback of a PV system. Therefore, several MPPT methods are proposed in order to extract maximum power from the PV module [1-3]. Reinforcement learning (RL) [4] is widely used in solving control problems since it can learn by interacting with the system without prior knowledge of the system model. In this paper, Q-learning [5] is used as the algorithm of the proposed methods and is implemented by constructing the Q-table (RL-QT MPPT). A solar MPPT method based on RL using a neural network with experience replay and fixed target Q-network technique (RL-QN MPPT) is also proposed, and it is expected to outperform the P&O method since the step size can be chosen according to the learned perturbation experience.

II. DESIGN OF AN REINFORCEMENT LEARNING MPPT SYSTEM

To perform MPPT based on RL, the system must be able to be described by Markov Decision Process (MDP). The element needed in RL MPPT system is defined in Table I. The goal of the agent is to reach the maximum power point (MPP) through interacting with the environment. The system’s condition, the state, is described by the solar irradiance, the module temperature and the duty ratio \( D \) since the \( I-V \) curve is affected by the solar irradiance and the module temperature as mentioned previously. The operating point of the system is at the intersection of the \( I-V \) curve and the load line, and the load line is controlled by the duty ratio of the converter. In this paper, the action is defined as a set of duty ratio changing step \( \Delta D \) with different step sizes. Therefore, the tracking progress can be seen as a sequential decision making problem, i.e., the MPP of the system can be reached by applying a series of variable step size \( \Delta D \) appropriately. The reward is a numerical signal that helps the agent judge the “good” or “bad” action. The action that moves the operating point close to the MPP is better than the action that moves the operating point away from the MPP. Therefore, the power difference \( \Delta P = P' - P \) is defined as the reward since it provides not only the moving direction of the operating point but also the numerical scaling representation of the effect caused by applying the action, for example, a larger step size may lead to a larger power difference. Also, the Markovian property holds since the current state is only affected by the state and the action taken one step before. Through combining the elements of the MDP model designed in Table I and the detail of the RL MPPT can be described as Fig.1. The perturbation step size is chosen according to the Q value of the current irradiance, temperature and duty ratio \( D \). After applying the change of \( D \), the power difference \( \Delta P \) and the new state description \( s' \) can be observed. In the Q-table approach, the experience will be used to update Q value immediately, but in the Q-network approach, it will be stored to perform experience replay.

| Parameter Needed to Perform RL | Parameter Selection in Solar MPPT System |
|-------------------------------|------------------------------------------|
| Environment                   | PV module and converter                  |
| Agent                         | controller                               |
| State                         | (irradiance, temperature, duty ratio)    |
| Action                        | \( \Delta D \)                            |
| Reward                        | \( \Delta P = P' - P \)                   |

Fig.1 Simple workflow of the RL MPPT
III. EXPERIMENT RESULT

The agent configuration is shown in Table II, which is identical for experiment. The corresponding hardware structure is depicted in Fig.3 and the actual experiment setup photos are presented in Fig.4. The experiments of the RL-QT MPPT method and the RL-QN MPPT method were conducted under similar environmental conditions. The irradiance was about 650 W/m², and the surface temperature of the PV module was about 48°C. The experiment results of the proposed system are shown in Fig.5 and Fig.6.

### TABLE II AGENT CONFIGURATION OF SIMULATION AND EXPERIMENT

|                  | RL-QT MPPT | RL-QN MPPT |
|------------------|------------|------------|
| **D range**      | 0.2-0.9    |            |
| **Sampling time**| 1s         |            |
| **ΔD (action)**  | [0, ±0.01, ±0.05, ±0.1] | | |
| **State**        | (irradiance, temperature, D) | (irradiance, temperature, D) |
| **Reward**       | ΔP         | ΔP         |
| **ε**            | 1          | 1          |
| **γ**            | 0.3        | 0.3        |
| **Q value storing type** | Q-table 4260*7 | Q-network 3-40-40-40-7 |
| **α**            | 0.9        | 0.0001     |
| **Cv**           | N/A        | 100        |
| **E**            | No limit   |            |
| **ph**           | 16         |            |

Fig.3 Hardware structure of the proposed system

Fig.4 The hardware setup with a 50 W, Voc=21.24V, Isc=3.05A PV module

Fig.5 P-t graph and D-t graph of the RL-QT MPPT experiment result

Fig.6 P-t graph and D-t graph of the RL-QN MPPT experiment result

IV. CONCLUSIONS

In this paper, two variable step size MPPT methods based on model free reinforcement learning are proposed. The tracking process can be seen as a sequential decision making problem since the MPP can be achieved through selecting appropriate perturbation step size for every time step. Therefore, an MDP model is suitable for describing the interaction between the circuit connected to the PV module and the controller which is able to choose ΔD and change the duty ratio D of the circuit.

An MDP model consists of four elements, which are state, action, transition and reward. With the MDP model described, an RL-QT MPPT method is proposed based on the Q-learning algorithm to perform MPPT control. However, the state representation is needed to be discretized for the tabular method, which may cause the loss of the MPPT control accuracy. Therefore, a Q-network approach is proposed. In RL-QN MPPT method, the table is approximated by a neural network, so that the discretization of the states are not needed. For both RL-QT MPPT and RL-QN MPPT, the tracking method consists of the learning phase and the tracking phase, which is able to expedite the tracking process since the Q value will not be updated in the tracking phase. A conclusion can be drawn from the result that the RL MPPT methods are effective methods to track the MPP since the variable step sizes are chosen appropriately to get close to the MPP faster and to remain stable near the MPP.

ACKNOWLEDGMENT

The authors would like to thank the Ministry of Science and Technology, Taiwan under Grant MOST 105-2511-S-009-016-MY3 for providing research funding.

REFERENCES

[1] A. Youssef, M. Telbany and A. Zekry, "Reinforcement Learning for Online Maximum Power Point Tracking Control", Journal of Clean Energy Technologies, vol. 4, no. 4, pp. 245-248, 2015.
[2] R. C. Hsu, C. T. Liu, W. Y. Chen, H. I. Hsieh and H. L. Wang, "A Reinforcement Learning-Based Maximum Power Point Tracking Method for Photovoltaic Array", International Journal of Photoenergy, vol. 2015, pp. 1-12, 2015.
[3] P. Kofinas, S. Doltsinis, A. Dounis and G. Vouros, "A reinforcement learning approach for MPPT control method of photovoltaic sources", Renewable Energy, vol. 108, pp. 461-473, 2017.
[4] R. S. Sutton and A. G. Barto, Reinforcement learning: An Introduction. MIT Press, 2018.
[5] C. Watkins and P. Dayan, "Q-learning", Machine Learning, vol. 8, pp. 257-277, 1992.