Reinforcement Learning for Online Maximum Power Point Tracking Control

Ayman Youssef, Mohamed El. Telbany, and Abdelhalim Zekry

***Abstract*—The world wide resource crisis led scientists and engineers to search for renewable energy sources. Photovoltaic systems are one of the most important renewable energy sources. In this paper we propose an intelligent solution for solving the maximum power point tracking problem in photovoltaic systems. The proposed controller is based on reinforcement learning techniques. The algorithm performance far exceeds the performance of traditional maximum power point tracking techniques. The algorithm not only reaches the optimum power it learns also from the environment without any prior knowledge or offline learning. The proposed control algorithm solves the problem of maximum power point tracking under different environment conditions and partial shading conditions. The simulations results show satisfactory dynamic and static response and superior performance over famous perturb and observe algorithm.**

***Index Terms*—Photovoltaic, maximum power point tracking, reinforcement learning.**

I. Introduction

Scientists all over the world are searching for renewable, clean, sustainable, and environmental energy sources. The sun is one of the most sustainable energy sources. Photovoltaic systems are responsible for converting the sun irradiance to electric power. The main problem with photovoltaic systems is its low efficiency which results in high cost of building such systems. There are many works in the literature to increase the efficiency of photovoltaic systems and hence decrease its cost. Artificial intelligence algorithms have been used in the design and control of photovoltaic systems. Their effectiveness in increasing the performance of photovoltaic systems led many researchers use them in designing, modeling, and controlling photovoltaic systems. In this work we introduce a reinforcement learning controller that controls the duty cycle of the boost DC/DC converter to insure maximum power is transferred from the PV panel.

The operation of maximum power point tracking is shown in Fig. 1. The controller reads the current and voltage from the solar panel and calculates the output duty cycle to control the DC/DC converter. There many work in the literature in designing and implementing the MPPT controller. AI techniques have proven their superiority in these algorithms.

In this work we introduce the reinforcement learning as a solution for the maximum power point tracking. This AI Algorithm learns from the environment and doesn't need prior learning or knowledge.

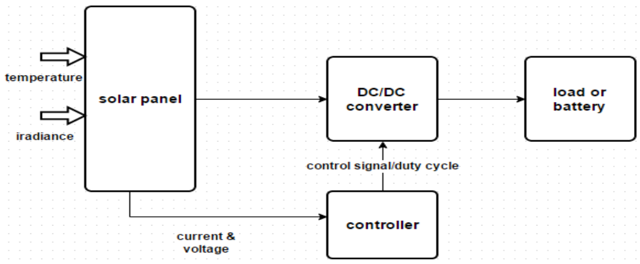

Fig. 1. MPPT.

II. MPPT for Photovoltaic Systems

In this section we introduce the model for the photovoltaic system and a survey on maximum power techniques introduced in the literature.

A. *Modeling Photovoltaic Panel*

The first step in building simulink models for photovoltaic systems is to model the photovoltaic panel. The photovoltaic panel is modeled as PV cells connected in series or parallel. The main focus of this research is maximum power point tracking techniques, so we use simple single diode model for each PV cell. The single diode model consists of a current source two resistors and one diode. The single diode model for photovoltaic systems is shown in Fig. 2.

The output of the current source is proportional to the solar energy incident on the solar cell. The photocell works as a diode when there is no incident solar energy. The shunt resistance models the internal losses due to current flow. Shunt resistance this due current leakage to the ground.

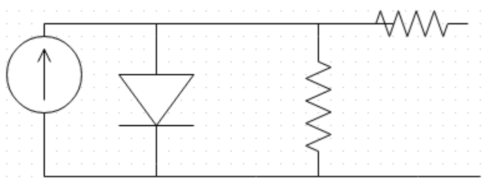
Fig. 2. photocell model.

The PV cell is modeled using the following equations first the thermal voltage *Vt*.

$$V_t = \frac{KTop}{q} \quad (1)$$

where K is Boltzmann's constant, Top is the operation temperature and q is the electron charge.

We need also to calculate the diode current:

Manuscript received May 2, 2015; revised August 12, 2015.
Ayman Youssef and Mohamed El. Telbany are with the Electronics Research Institute, Giza, Egypt (e-mail: aymanmahgoub@eri.sci.eg, telbany@eri.sci.eg).
Abdelhalim Zekry is with the Electronics and Communication Department, Ain Shams University, Egypt (e-mail: author@nrim.go.jp).

$$I_D = \left[e^{\frac{V+IRs}{nV_tCN_s}} - 1 \right] I_S N_S \quad (2)$$

where n is ideality factor, N_s number of series cells, C is the number of cells in each module and I_S means the diode reverse saturation current.

I_s is calculated by using the following equation:

$$I_S = I_{rs} \left(\frac{Tc}{Tref}\right)^3 e^{\left[\frac{qEg}{nk}\left(\frac{1}{Tc} - \frac{1}{Tref}\right)\right]} \quad (3)$$

where Eg is the band gap and $Tref$ is the reference temperature

$$I_{rs} = \frac{I_{sc}}{\left[e^{\frac{Vocq}{kCTcn}} - 1 \right]} \quad (4)$$

where Voc open circuit voltage and Isc is the short circuit current. Voc, Isc can be determined from the PV array that is being modeled.

$$I_{sh} = \frac{V + IRs}{Rsh} \quad (5)$$

$$I_{ph} = Gk\left[Isc + Ki(Tc - Tref)\right] \quad (6)$$

$$I = NpIph - Id - Ish \quad (7)$$

Fig. 3 shows the power VS voltage plot with different irradiance values.

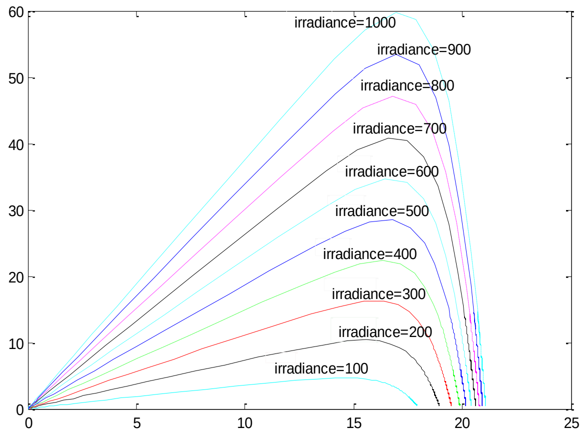

Fig. 3. Power VS voltage.

Fig. 3 shows the effectiveness of the photovoltaic single diode model. Fig. 3 shows that we have different maximum power point for different irradiance values.

B. Review of Maximum Power Point Techniques

There many techniques in the literature for maximum power point tracking. Perturb and observe is the most implemented algorithm for MPPT due its simplicity [1], [2]. The algorithm however suffers from oscillations around the maximum power point especially under varying atmospheric conditions. This results in power loss and reduced system efficiency. Another algorithm that is always discussed in literature is the increment conductance algorithm.

AI algorithms on the other hand have proven their superiority over traditional algorithms. AI algorithms can reach the maximum power point with high accuracy and speed saving power losses of the system.

The most wide spread AI technique for maximum power point tracking is the fuzzy controller [3]-[5]. There are many works in the literature for fuzzy controller design. Ant colony is used to design the membership functions for the fuzzy controller [6].

There are many work also done in maximum power point tracking using neural networks.

The main drawback of fuzzy controllers is the need for rules and membership functions to represent the knowledge. In this work we propose using reinforcement learning controller. The controller learns from the environment by exploring it. The algorithm hence doesn't require complex offline learning algorithms.

III. Units

In this work we introduce a reinforcement learning solution for the maximum power point tracking. Reinforcement learning algorithms have the ability of learning from the environment on-line. This is a quite advantage as photovoltaic systems operate under different climate and weather conditions. In this work we propose a temporal difference Q-learning algorithm.

A. The Reinforcement Learning Model

The general model of a reinforcement learning algorithm includes agent, environment, state, action, and reward. The agent explores the environment and the appropriate actions to achieve a predefined goal. The agent learns from environment through the reward function. The agent maintains a reward average value for a certain state-action pair. When the agent finishes exploration it starts the exploitation phase.

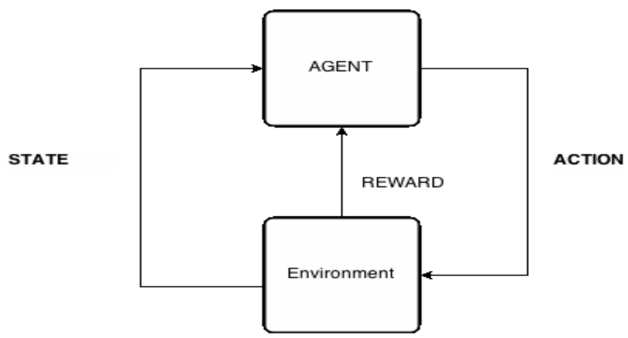

Fig. 4. Reinforcement learning.

B. Formulation of the Q-learning Model on MPPT

The reinforcement learning algorithm consists of states, actions, and reward. In this section, we define the three concepts for our MPPT problem.

1) States

In the formulation of the maximum power point we have four states. S= [$S1$, $S2$, $S3$, $S4$]. $S1$ is the state of going towards the maximum power point from the right. $S2$ is the state going towards the maximum power point from the right. $S3$ is the state leaving the maximum power point from the left. $S4$ is the state leaving the maximum power point from the left.

2) *Actions*

In our formulation of the maximum power point tracking we have four actions. These actions represent the step decrease or increase of the duty cycle. We start the duty cycle of the boost inverter switch with .5. The controller then decides the step to decrease or increase the duty cycle of the pulse width generated signal. We have four actions decreasing or increasing the duty cycle by .01, and decreasing or increasing the duty cycle by .05. We have four action A = [-.005 -.01 .005 .01].

3) *Reward*

In the reinforcement learning algorithm we need a definition for the reward function. In our work we give the reward a positive value if the power increases and a negative value when the power decreases. If the power change is small the reward function returns zero.

In this work we introduce a q-learning algorithm. The Q is a matrix that assigns an action to a given state.

$$Q(s,a) = Q(s,a) + \eta \Delta Q(s,a) \tag{8}$$

$$\Delta Q(s,a) = \left[r + \gamma Qm(sn) \right] - Q(s,a) \tag{9}$$

$$Qm(sn) = \max_{\text{all } b \text{ from } A} Q(sn,b) \tag{10}$$

where s represents the state, a represents the action assigned by the controller to this state, and sn represents the new state. The learning algorithm searches for the optimum value for the Q matrix to reach optimum control. The algorithm updates the q matrix values depending on the reward value.

IV. Simulink Results and Dissection

A. *Maximum Power Point Tracking*

To implement the design we used a MATLAB coding function. Inside this block we wrote the code for the q-learning algorithm.

Fig. 6 shows the input irradiance variations. To test the speed of the control we apply step changes to the input irradiance and monitor the photocell output power.

Fig. 5 shows the output power generated from the system using perturb and observe algorithm vs. reinforcement learning algorithm. It is obvious that the q-learning algorithm has lower oscillations and higher speed than perturb and observe. The most important advantage of the reinforcement learning algorithms is that it learns online from the environment. Other AI algorithms like fuzzy or neural networks need offline learning with complex techniques and large data sets. The reinforcement learning learns from environment and this makes it suitable for photovoltaic systems under varying atmospheric conditions.

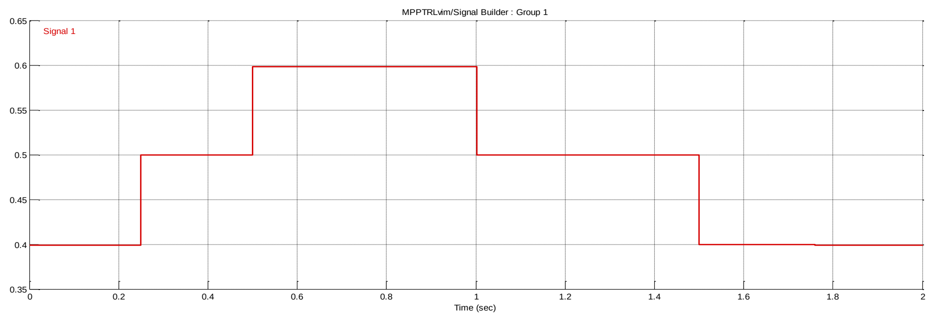

Fig. 5. Input irradiance.

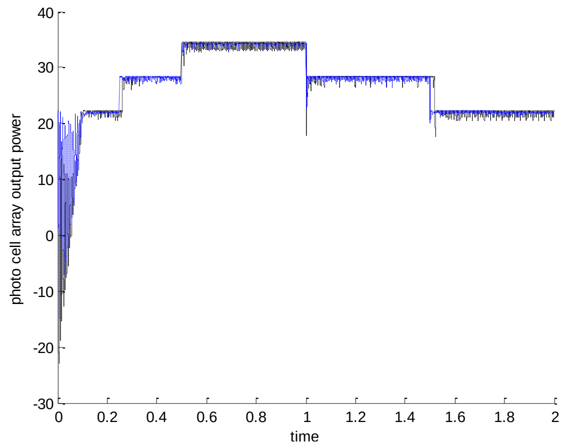

Fig. 6. Photovoltaic power P&O vs RL.

Fig. 6 shows the duty cycle output from the P&O algorithm and the reinforcement learning algorithm. It can be seen from the graph that both controllers reach the same duty cycle. The reinforcement learning is faster than the P&O algorithm and has fewer oscillations.

V. Conclusion

This work introduces a reinforcement learning algorithm for the maximum power point tracking problem. The proposed AI algorithm learns directly from the environment and has higher dynamic speed than P&O and lower oscillations.

References

[1] J. J. Nedumgatt, K. B. Jayakrishnan, S. Umashankar, D. Vijayakumar, and D. P. Kothari, “Perturb and observe MPPT algorithm for solar PV systems-modeling and simulation,” in *Proc. Annu. IEEE India Conf.*, pp. 1-6.

[2] L. Piegari and R. Rizzo, “Adaptive perturb and observe algorithm for photovoltaic maximum power point tracking,” *IET Renewable Power Gener.*, vol. 4, no. 4, pp. 317-328.

[3] T. L. Kottas, Y. S. Boutalis, and A. D. Karlis, "New maximum power Point tracker for PV arrays using fuzzy controller in close cooperation with fuzzy cognitive networks," *IEEE Transaction on Energy Conversion*, vol. 21, no. 3, pp. 793–803, September 2006.

[4] J. Li and H. Wang, "Maximum power point tracking of photovoltaic generation based on the fuzzy control method," in *Proc. SUPERGEN'09*, April 2009, pp. 1-6, 6-7.

[5] D. Menniti, A. Pinnarelli, and G. Brusco, "Implementation of a novel fuzzy logic based MPPT for grid-connected photovoltaic generation system," in *Proc. Trondheim PowerTech.*, June 19-23, 2011, pp. 1–7.

[6] C. Larbes, S. M. A. Cheikh, T. Obeidi, A. Zerguerras, "Genetic algorithms optimized fuzzy logic control for the maximum power point racking in photovoltaic system," *Renewable Energy*, vol. 34, issue 10, 2093–2100, 2009.

Ayman Youssef was born in Iowa State, on April 16, 1986. The author has a bachelor degree from Electronics and Communication Department, Cairo University in 2008. He earned his master degree from Cairo University, Electronics and Communication Department in 2012. He is currently working toward his PhD degree in photovoltaic systems in Ain Shams University.

He worked as a teaching assistant in Akhbar El Youm Academy. He now works as a research assistant in Electronics Research Institute, Giza. He also works as an instructor in private sector universities and institutes. He has 6 publications in neural network FPGA implementations and photovoltaic systems. He is interested in FPGA design, artificial intelligence, image processing, photovoltaic systems, and classification. He also worked as a reviewer in 2012 IEEE International Conference on Electronics Design, Systems and Applications Conference.

Mohamed El. Telbany was born in Dammita, on August 21, 1968. He has earned a bachelor degree from the Computer Engineering Department, Minoufia University in 1991. he earned a master degree from the Faculty of Engineering, Cairo University in 1997. He earned a PhD degree in computers engineering, from the Department of Electronics and Communications in 2003.

The author worked as an instructor in many universitieis. He worked as an associate professor in Islamic University in Madina, Taif University, and King Khaled University. He is currently working as an associate professor in Electronics Research Institute. He also worked as an instructor and gave many courses. His research interests include artificial intelligence, particle swarm optimization, forcasting, robotics, and neural networks,

Dr. Mohamed El. Telbany was awarded the United Nations Development Program (UNDP) and European Space Agency research fellowship in Remote Sensing and Information Technology.

Abdellhalim Zekry was born in Menofia on August 8, 1946. He earned his bachelor degree in electrical engineering from Cairo University in 1969. He earned his master degree from the Electrical Engineering Department, Cairo University. The author earned his PhD degree from Technical University of Berlin, West Germany 1981.

He worked as a staff member in many universities. He is a member in IEEE society. He is also a member in Egyptian Physical Society. He made many consolations for governmental private sectors in Egypt and Saudi Arabia. He taught many courses of electronics at Cairo University, TU Berlin, Ain shams university, and king Saudi university.

Prof. Abdellhalim Zekry was awarded Ain Shams University Research Prize, he earned the best paper award in China in 2011. He was included in who is who in the world several times for achievements. He earned the Egypt Encouragment State prize in engineering sciences in 1992. He also earned the recognition prize of Ain Shams University in technological sciences in 2010.

Bioenergy