Q-Learning Fuzzy PID Controller Design for Motor Control

Ching-Chang Wong, I-Shen Yeh, and Shao-Yu Chien

***Abstract*—In this paper, a fuzzy PID control method based on Q-learning is proposed to control the motor so that it can adapt to different environments and meet the expectations of the request. There are two main parts: (1) a fuzzy control method is proposed to adjust the parameters K_P, K_I, and K_D of the PID controller and (2) a Q-learning algorithm is proposed to learn the fuzzy rule base and the membership functions of the fuzzy variable. The fuzzy method is proposed to modify the parameters K_P, K_I, and K_D of the PID controller, where the K_P, K_I, and K_D of the PID controller will be automatically adjusted according to environmental changes or external disturbance. The Q-learning algorithm is proposed to learn the fuzzy rule base and the membership functions of the fuzzy variables. The Q-learning algorithm lets the fuzzy rule base and the membership functions of the fuzzy variable that originally relied on the expert rule can be obtained through repeated learning. A sliding mode is added in the Q-learning algorithm and fuzzy control to reduce the number of system parameters required in the learning process to improve the learning efficiency. The learning process is to learn membership functions of the fuzzy variables with the initial fuzzy rule base and the initial membership functions of the fuzzy variable, and then learn a new fuzzy rule base with the initial fuzzy rule base and the new membership functions of the fuzzy variable. Some experimental results of the voice coil motor, brushed DC motor, and Brushless DC motor are presented to illustrate that the proposed method can indeed effectively control these three motors.**

Index Terms*—*Fuzzy Control, Sliding Mode, PID Controller, Motor Control, Q-learning

I. INTRODUCTION

Automation equipment is widely used in industrial manufacturing today, and these equipment require power to complete their actions. Motors are the most widely used power source, and current industrial development still focuses on motors as the main power output. There are many types of motors. For Voice Coil Motor (VCM), it is suitable for short-stroke linear motion, high speed, and high acceleration/deceleration application environment. For brushed DC motor, the control conditions are relatively simple. It only needs to control the magnitude and polarity of the voltage to change the speed and direction of the motor. For Brushless DC (BLDC) motors, a sequential control voltage phase is required to change the speed and direction of the motor, but compared with the brushed DC motor, there is no need to consider the problem of the loss of the electric brush, so it has a longer service life. After installing the encoder on the motor as the feedback signal of the position or speed, and using the controller to approach the feedback value to the control target value for closed loop control, it can be called a servo motor. As far as the industry is concerned, most of its control methods still use Proportional Integral Derivative (PID) controllers, which are characterized by simple system architecture and easy implementation, and have good performance for most types of systems. When the traditional PID controller [1] controls a highly complex system, its parameters are not easy to adjust, which makes it impossible to provide sufficient performance for the system. After Mamdani [2] carried out the first application of fuzzy control, fuzzy control became an alternative method [3]. Fuzzy control has the advantage of expert experience, and can be combined with traditional controllers to handle complex control systems. Various types of fuzzy PID controllers have been proposed. According to the application of fuzzy theory in PID controllers and their control architecture [4], they can be divided into two categories: (1) PID-type fuzzy controllers [5][6][7], and (2) fuzzy methods are used to tune the gain parameters of the PID controller or they are self-adaptive [8][9]. The structure of the PID-type fuzzy controller is similar to that of the conventional PID controller. It can be achieved by combining PI-type and PD-type fuzzy controllers with two different fuzzy rule bases, or combining a PD-type fuzzy controller with an integrator to achieve. The structure in which the gain parameters of the PID controller are automatically adjusted by a fuzzy method is to adjust the gain parameters of the traditional PID controller online by a fuzzy method, and the PID controller outputs control signals. In fuzzy control, the main adjustable parts are divided into two categories: (1) structure tuning, and (2) parameter tuning. The part related to the adjustment of structure is the structure of rules, the number of rules, the meaning of variables, and the division of each variable area. The part related to the adjustment of the parameters is the shape and position of membership functions, such as the adjustment of the center and width of the triangular membership function. For the automatic learning or adjusting the fuzzy control system, unsupervised learning can grasp the regularity of the input vector and construct the model without external information, so it is suitable for clustering of data and find out the corresponding rule. Supervised learning and Reinforcement Learning (RL) [10][11] are usually used to adjust rules or membership functions of the fuzzy system. Genetic Algorithm (GA) can be used to make structural adjustments, the method used is to search for all possible fuzzy rules in the space [12][13][14].

There have been many studies on fuzzy PID controllers and adaptive fuzzy controllers. In order to propose a method that conforms to the motor control, this paper will combine the above two methods and use the Q-learning algorithm in reinforcement learning [15][16][17] to adjust the structure and parameters of the fuzzy controller and make the parameters of the PID controller have self-adaptability.

This paper was first submitted on November 30, 2020. This work was supported in part by the Ministry of Science and Technology (MOST), Taiwan, under Grant MOST 108-2221-E-032-045 and MOST 109-2221-E-032-038.

Ching-Chang Wong, I-Shen Yeh, and Shao-Yu Chien are with the Department of Electrical and Computer Engineering, Tamkang University, New Taipei City, Taiwan.

Corresponding author: Ching-Chang Wong (e-mail: wong@ee.tku.edu.tw)

II. Q-LEARNING FUZZY PID CONTROL ARCHITECTURE

The Q-learning fuzzy PID controller proposed in this paper mainly uses the fuzzy method to adjust the parameters K_P, K_I and K_D of the PID controller to make it adaptive. Then the Q-learning algorithm is added, so that the fuzzy rule base and the membership functions of fuzzy variables that originally relied on expert rules can be obtained through repeated learning, and the sliding mode is added to the Q-learning algorithm and fuzzy control[18][19], in order to reduce the number of system parameters required in the learning process, thereby improving the efficiency of learning. There are two main parts: (1) adjust the PID controller by the fuzzy method and (2) learn the fuzzy rule base and the membership functions of the fuzzy variables by the Q-learning algorithm.

A. Adjust the PID Controller by the Fuzzy Method

The fuzzy PID controller in this paper is used to adjust the K_P, K_I, and K_D in the PID controller according to the error e and the error change Δe of the motor input command and feedback, with the membership functions and the fuzzy rule base in the controller. After the error e is input to the PID controller, the PID controller calculates the output motor command. Therefore, the fuzzy PID controller can adjust the appropriate K_P, K_I and K_D parameters according to environmental changes or external disturbance. The control diagram is shown in Fig. 1.

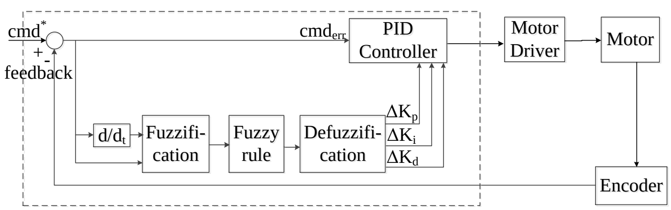

Fig. 1. Fuzzy PID control architecture.

The selection of the domain of the fuzzy controller's input variables e and Δe and output variables ΔK_P, ΔK_I, and ΔK_D are defined as follows:

$$e \in [-80, 80] \tag{1}$$

$$\Delta e \in [-6, 6] \tag{2}$$

$$\Delta K_P \in [-0.8K_P, 0.8K_P] \tag{3}$$

$$\Delta K_I \in [-0.4K_I, 0.4K_I] \tag{4}$$

$$\Delta K_D \in [-0.8K_D, 0.8K_D] \tag{5}$$

In the selection of fuzzy linguistic items, five items are defined in the respective domains of the input variables e and Δe and the output variables ΔK_P, ΔK_I, and ΔK_D, which are expressed as follows:

$$T(e) = \{NB, NS, ZO, PS, PB\} = \{A_0, A_1, A_2, A_3, A_4\} \tag{6}$$

$$T(\Delta e) = \{NB, NS, ZO, PS, PB\} = \{B_0, B_1, B_2, B_3, B_4\} \tag{7}$$

$$T(\Delta K_P) = \{NB, NS, ZO, PS, PB\} = \{C_0, C_1, C_2, C_3, C_4\} \tag{8}$$

$$T(\Delta K_I) = \{NB, NS, ZO, PS, PB\} = \{D_0, D_1, D_2, D_3, D_4\} \tag{9}$$

$$T(\Delta K_D) = \{NB, NS, ZO, PS, PB\} = \{E_0, E_1, E_2, E_3, E_4\} \tag{10}$$

where NB, NS, ZO, PS, and PB are respectively used to denote Negative Big, Negative Small, Zero, Positive Small, and Positive Big.

The triangular membership function is used to describe the fuzzy set of input variables e and Δe. As shown in Fig. 2 and Fig. 3, $\mu_A(e)$ and $\mu_B(\Delta e)$ represent the degrees of certainty of these two input variables, respectively. In the definition of the fuzzy sets of output variables ΔK_P, ΔK_I, and ΔK_D, the fuzzy singleton is used to describe the fuzzy set of output variables. As shown in Fig. 4, Fig. 5, and Fig. 6,$\mu_C(\Delta K_P)$, $\mu_D(\Delta K_I)$, and $\mu_E(\Delta K_D)$ are their respective degrees of certainty.

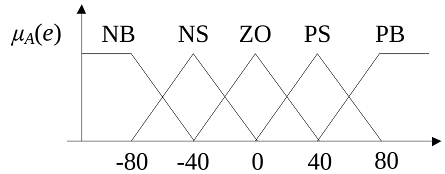

Fig. 2. Membership functions of the fuzzy input variable e.

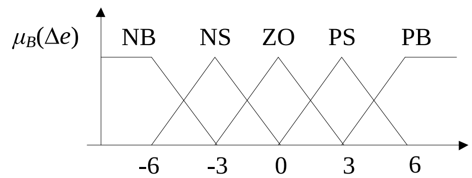

Fig. 3. Membership functions of the fuzzy input variable Δe.

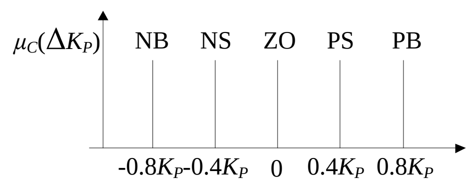

Fig. 4. Membership functions of the fuzzy input variable ΔK_P.

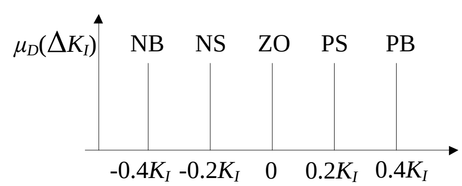

Fig. 5. Membership functions of the fuzzy output variable ΔK_I.

Fig. 6. Membership functions of the fuzzy output variable ΔK_D.

In the establishment of the fuzzy rule base, the fuzzy rule base of ΔK_P, ΔK_I, and ΔK_D constructed in this paper are respectively shown in Table I, Table II, and Table III.

TABLE I
FUZZY RULE BASE OF ΔK_P

ΔK_P		e				
		NB(A_0)	NS(A_1)	ZO(A_2)	PS(A_3)	PB(A_4)
Δe	NB(B_0)	PB(C_4)	PS(C_3)	PS(C_3)	PS(C_3)	ZO(C_2)
	NS(B_1)	NS(C_1)	NS(C_1)	NS(C_1)	ZO(C_2)	NS(C_1)
	ZO(B_2)	NS(C_1)	NS(C_1)	ZO(C_2)	PS(C_3)	NS(C_1)
	PS(B_3)	NS(C_1)	ZO(C_2)	PS(C_3)	PS(C_3)	NS(C_1)
	PB(B_4)	ZO(C_2)	PS(C_3)	PS(C_3)	PB(C_4)	NB(C_0)

TABLE II
FUZZY RULE BASE OF ΔK_I

ΔK_I		e				
		NB(A_0)	NS(A_1)	ZO(A_2)	PS(A_3)	PB(A_4)
Δe	NB(B_0)	NB(D_0)	NB(D_0)	NS(D_1)	NS(D_1)	ZO(D_2)
	NS(B_1)	NS(D_1)	NS(D_1)	NS(D_1)	ZO(D_2)	PS(D_3)
	ZO(B_2)	NS(D_1)	NS(D_1)	ZO(D_2)	PS(D_3)	PS(D_3)
	PS(B_3)	NS(D_1)	ZO(D_2)	PS(D_3)	PS(D_3)	PS(D_3)
	PB(B_4)	ZO(D_2)	PS(D_3)	PS(D_3)	PB(D_4)	PB(D_4)

TABLE III
FUZZY RULE BASE OF ΔK_D

ΔK_D		e				
		NB(A_0)	NS(A_1)	ZO(A_2)	PS(A_3)	PB(A_4)
Δe	NB(B_0)	PS(E_3)	ZO(E_2)	ZO(E_2)	ZO(E_2)	PB(E_4)
	NS(B_1)	NB(E_0)	NS(E_1)	NS(E_1)	ZO(E_2)	PS(E_3)
	ZO(B_2)	NB(E_0)	NS(E_1)	NS(E_1)	ZO(E_2)	PS(E_3)
	PS(B_3)	NB(E_0)	NS(E_1)	NS(E_1)	ZO(E_2)	PS(E_3)
	PB(B_4)	PS(E_3)	ZO(E_2)	ZO(E_2)	ZO(E_2)	PB(E_4)

In the selection of fuzzification interface and defuzzification method, the Sugeno's fuzzy inference method and weighted average method is adopted so that the output of the fuzzy system can be expressed by

$$v_m = \frac{\sum_{j_1=1}^{5}\sum_{j_2=1}^{5} w(j_1, j_2)\cdot y}{\sum_{j_1=1}^{5}\sum_{j_2=1}^{5} w(j_1, j_2)} \tag{11}$$

where $y \in \{C_{j3}, D_{j3}, E_{j3}\}$ is the definite value represented by the fuzzy single point output.

B. Learn the Fuzzy Rule Base and Membership Functions of Fuzzy Variables by the Q-learning Algorithm

The Q-learning fuzzy PID controller proposed in this paper mainly adds the Q-learning algorithm, so that the fuzzy rule base and the membership functions of fuzzy variables that originally relied on expert rules can be obtained through repeated learning. A sliding mode is added to the Q-learning algorithm and fuzzy control to reduce the number of system parameters required in the learning process, thereby improving the efficiency of learning. The control architecture diagram is shown in Fig. 7. The learning process is to first learn the membership functions of the new fuzzy variable from the initial fuzzy rule base and the membership functions of the fuzzy variable, and then learn the new fuzzy rule from the initial fuzzy rule base and the membership functions of the new fuzzy rule base. The design architecture can be divided into three main projects: 1) establishing a sliding plane, 2) using Q-learning to learn and adjust membership functions of fuzzy variables, and 3) using Q-learning to compare the fuzzy rule base for learning and adjustment.

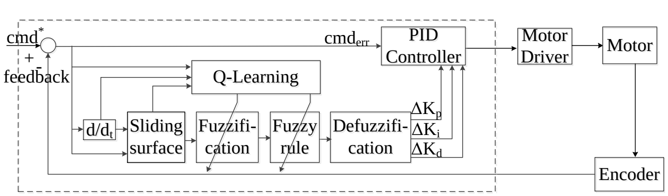

Fig. 7. Q-learning fuzzy PID control architecture

1) Establishing a Sliding Plane

In the original fuzzy PID, the input of the fuzzy controller is 2 inputs, which are the error e and the error change Δe, so that when Q-learning is learning and adjusting the fuzzy controller, there are 28 parameters in the part of the membership functions that need to be learned and adjusted. There are 375 parameters in the fuzzy rule base that need to be learned and adjusted. After the sliding plane is established, the error e and the error change Δe can be processed, as shown in equation (12):

$$s = \Delta e + \lambda e \tag{12}$$

where λ is the expansion parameter and s is the output of the sliding plane. The output s obtained after sliding plane processing can be used as the input of the fuzzy controller, so that when Q-learning learns and adjusts the fuzzy controller, there are only 14 parameters in the membership functions that need to be learned and adjusted. In the part of the fuzzy rule base, only 75 parameters need to be learned and adjusted, which greatly improves the efficiency of Q-learning in learning and adjusting the fuzzy controller.

2) Using Q-Learning to Learn and Adjust Membership Functions of Fuzzy Variables

First define the membership functions of the initial fuzzy variable of the fuzzy controller and the fuzzy rule base. The membership functions of the fuzzy variable is equally distributed based on the value range of the sliding plane output s. The membership functions of the fuzzy variable are shown in Fig. 8. The fuzzy rule base is based on the common ΔK_P, ΔK_I, and ΔK_D fuzzy rule base designed with the error e and the error change Δe in the references, and the fuzzy rule of its symmetry axis is taken as the fuzzy rule of single input s. The rule bases are shown in Table IV, Table V, and Table VI.

TABLE IV
FUZZY RULE BASE OF ΔK_P FOR SINGLE INPUT S

ΔK_P					
s	NB	NS	ZO	PS	PB
u	PB	NS	ZO	PS	NB

TABLE V
FUZZY RULE BASE OF ΔK_I FOR SINGLE INPUT S

ΔK_I					
s	NB	NS	ZO	PS	PB
u	NS	NB	ZO	PS	PB

TABLE VI
FUZZY RULE BASE OF ΔK_D FOR SINGLE INPUT S

ΔK_D					
s	NB	NS	ZO	PS	PB
u	PS	NS	NS	ZO	PB

Then define the states and actions of Q-learning to build the Q table. The input of the fuzzy controller is the state, and the adjustment value of the membership function is the action. The relationship between the state and the input variable sliding plane output s is shown in Table VII, and the relationship between the adjustment value α and β of the action and the membership function is shown in Fig. 8. The Q table for Q-learning to learn the membership function is shown in Table VIII.

TABLE VI
THE RELATIONSHIP BETWEEN THE STATE AND THE INPUT VARIABLE SLIDING PLANE OUTPUT S

when	$s < -60$	*then*	*state = 2*
when	$-60 \le s < -20$	*then*	*state = 1*
when	$-20 \le s < 20$	*then*	*state = none*
when	$20 \le s < 60$	*then*	*state = 1*
when	$s \ge 60$	*then*	*state = 2*

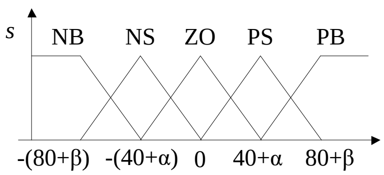

Fig. 8. The relationship between the adjustment value α and β of the action and the membership function.

TABLE VIII
THE Q TABLE FOR Q-LEARNING TO LEARN MEMBERSHIP FUNCTIONS

Q-Table							
State	Action						
s	-9	-6	-3	0	3	6	9
1	α_0	α_1	α_2	α_3	α_4	α_5	α_6
2	β_0	β_1	β_2	β_3	β_4	β_5	β_6
none	0	0	0	0	0	0	0

After setting the state, action and Q table, we have to defined the reward function of Q-learning. This paper designs the reward function with the error e and the error change Δe. The goal is to hope that the error e is small and the error change Δe is large. The reward function is defined by

$$R(e,\Delta e)=\frac{1}{1+|e|}-\frac{1}{1+|\Delta e|} \tag{13}$$

3) Using Q-learning to Compare the Fuzzy Rule Base for Learning and Adjustment

Use membership functions of the adjusted fuzzy variable and the defined fuzzy rule base of single input *s* as the initial value to learn and adjust the fuzzy rule base, then define the state and action of Q-learning to establish Q-table. The input of the fuzzy controller is the state, the fuzzy rule adjustment value of ΔK_P, ΔK_I, and ΔK_D belongs to the action. The relationship table between the state and the input variable sliding plane output s is shown in Table IX, the action and fuzzy rule base adjustment value γ_P, γ_I, γ_D relation table shown in Table X, Table XI, and Table XII. The Q-table for Q-learning to learn the fuzzy rule base are shown in Table XIII, Table XIV, and Table XV.

TABLE IX
THE RELATIONSHIP TABLE BETWEEN THE STATE AND THE INPUT VARIABLE SLIDING PLANE OUTPUT S

when	$s < -60$	*then*	*state = 1*
when	$-60 \le s < -20$	*then*	*state = 2*
when	$-20 \le s < 20$	*then*	*state = 3*
when	$20 \le s < 60$	*then*	*state = 4*
when	$s \ge 60$	*then*	*state = 5*

TABLE X
THE ACTION AND ΔK_P FUZZY RULE BASE ADJUSTMENT VALUE Γ_P RELATION TABLE

ΔK_P					
s	NB	NS	ZO	PS	PB
u	PB or $\gamma_{P(0,n)}$	NS or $\gamma_{P(1,n)}$	ZO or $\gamma_{P(2,n)}$	PS or $\gamma_{P(3,n)}$	NB or $\gamma_{P(4,n)}$

TABLE XI
THE ACTION AND ΔK_I FUZZY RULE BASE ADJUSTMENT VALUE Γ_I RELATION TABLE

ΔK_I					
s	NB	NS	ZO	PS	PB
u	NS or $\gamma_{I(0,n)}$	NO or $\gamma_{I(1,n)}$	ZO or $\gamma_{I(2,n)}$	PS or $\gamma_{I(3,n)}$	PB or $\gamma_{I(4,n)}$

TABLE XII
THE ACTION AND ΔK_D FUZZY RULE BASE ADJUSTMENT VALUE Γ_D RELATION TABLE

ΔK_I					
s	NB	NS	ZO	PS	PB
u	PS or $\gamma_{D(0,n)}$	NS or $\gamma_{D(1,n)}$	NS or $\gamma_{D(2,n)}$	ZO or $\gamma_{D(3,n)}$	PB or $\gamma_{D(4,n)}$

TABLE XIII
THE Q-TABLE FOR Q-LEARNING TO LEARN THE ΔK_P FUZZY RULE BASE

ΔK_P					
State	Action				
s	NB	NS	ZO	PS	PB
1	$\gamma_{P(0,0)}$	$\gamma_{P(0,1)}$	$\gamma_{P(0,2)}$	$\gamma_{P(0,3)}$	$\gamma_{P(0,4)}$
2	$\gamma_{P(1,0)}$	$\gamma_{P(1,1)}$	$\gamma_{P(1,2)}$	$\gamma_{P(1,3)}$	$\gamma_{P(1,4)}$
3	$\gamma_{P(2,0)}$	$\gamma_{P(2,1)}$	$\gamma_{P(2,2)}$	$\gamma_{P(2,3)}$	$\gamma_{P(2,4)}$
4	$\gamma_{P(3,0)}$	$\gamma_{P(3,1)}$	$\gamma_{P(3,2)}$	$\gamma_{P(3,3)}$	$\gamma_{P(3,4)}$
5	$\gamma_{P(4,0)}$	$\gamma_{P(4,1)}$	$\gamma_{P(4,2)}$	$\gamma_{P(4,3)}$	$\gamma_{P(4,4)}$

TABLE XIV
THE Q-TABLE FOR Q-LEARNING TO LEARN THE ΔK_I FUZZY RULE BASE

ΔK_I					
State	Action				
s	NB	NS	ZO	PS	PB
1	$\gamma_{I(0,0)}$	$\gamma_{I(0,1)}$	$\gamma_{I(0,2)}$	$\gamma_{I(0,3)}$	$\gamma_{I(0,4)}$
2	$\gamma_{I(1,0)}$	$\gamma_{I(1,1)}$	$\gamma_{I(1,2)}$	$\gamma_{I(1,3)}$	$\gamma_{I(1,4)}$
3	$\gamma_{I(2,0)}$	$\gamma_{I(2,1)}$	$\gamma_{I(2,2)}$	$\gamma_{I(2,3)}$	$\gamma_{I(2,4)}$
4	$\gamma_{I(3,0)}$	$\gamma_{I(3,1)}$	$\gamma_{I(3,2)}$	$\gamma_{I(3,3)}$	$\gamma_{I(3,4)}$
5	$\gamma_{I(4,0)}$	$\gamma_{I(4,1)}$	$\gamma_{I(4,2)}$	$\gamma_{I(4,3)}$	$\gamma_{I(4,4)}$

TABLE XV
THE Q-TABLE FOR Q-LEARNING TO LEARN THE ΔK_D FUZZY RULE BASE

ΔK_D					
State	Action				
s	NB	NS	ZO	PS	PB
1	$\gamma_{D(0,0)}$	$\gamma_{D(0,1)}$	$\gamma_{D(0,2)}$	$\gamma_{D(0,3)}$	$\gamma_{D(0,4)}$
2	$\gamma_{D(1,0)}$	$\gamma_{D(1,1)}$	$\gamma_{D(1,2)}$	$\gamma_{D(1,3)}$	$\gamma_{D(1,4)}$
3	$\gamma_{D(2,0)}$	$\gamma_{D(2,1)}$	$\gamma_{D(2,2)}$	$\gamma_{D(2,3)}$	$\gamma_{D(2,4)}$
4	$\gamma_{D(3,0)}$	$\gamma_{D(3,1)}$	$\gamma_{D(3,2)}$	$\gamma_{D(3,3)}$	$\gamma_{D(3,4)}$
5	$\gamma_{D(4,0)}$	$\gamma_{D(4,1)}$	$\gamma_{D(4,2)}$	$\gamma_{D(4,3)}$	$\gamma_{D(4,4)}$

In this paper, the reward function is designed based on the error e and the error change Δe. The goal is to hope that the error e is small and the error change Δe is large. The reward function is defined by

$$R(e,\Delta e)=\frac{1}{1+|e|}-\frac{1}{1+|\Delta e|} \tag{14}$$

In the control architecture of this paper, the motor position control of the voice coil motor experimental platform will be discussed. The physical diagram is shown in Fig. 9, and the control architecture diagram is shown in Fig. 10. The physical diagram of the Brushless DC motor experimental platform of motor speed control is shown in Fig. 11, and the control architecture diagram is shown in Fig. 12. The motor speed control of the brushed DC motor of the two-wheel self-balancing vehicle is shown in Fig. 13, and the control architecture diagram is shown in Fig. 14.

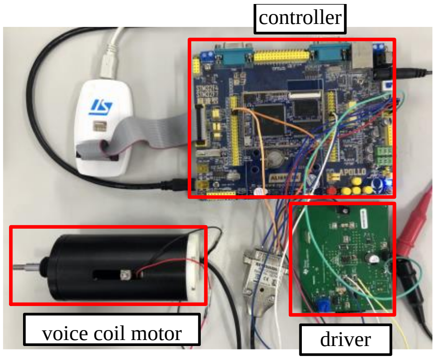

Fig. 9. Physical diagram of the voice coil motor experimental platform.

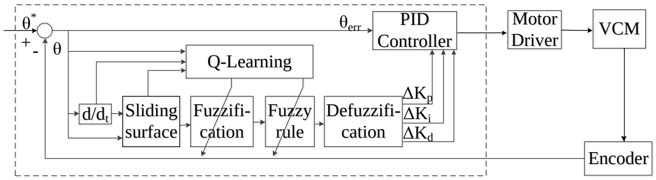

Fig. 10. Position control structure of voice coil motor.

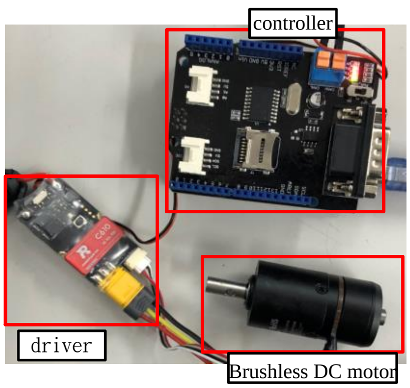

Fig. 11. Physical diagram of the Brushless DC motor experimental platform.

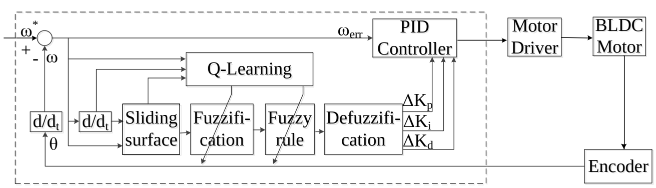

Fig. 12. Speed control architecture of Brushless DC motor.

Fig. 13. Physical diagram of two-wheeled self-balancing vehicle.

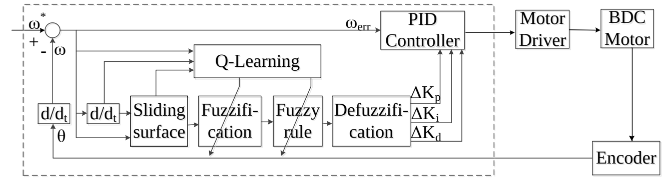

Fig. 14. Speed control structure of brushed DC motor of two-wheel self-balancing vehicle.

III. Experimental Results

A. Position Control of the Voice Coil Motor Experimental Platform

In the process of learning and adjusting the membership functions of the fuzzy variable of the fuzzy controller by Q-learning, it is mainly to sequentially give the motor of the voice coil motor experiment platform 100% position command and 0% position command, a total of 2 position commands. Then repeat these two commands 100 times in sequence. The membership functions of the fuzzy variables before and after learning are shown in Fig. 15. The fuzzy rule base of ΔK_P, ΔK_I, and ΔK_D before and after learning are shown in Table XVI, Table XVII, and Table XVIII.

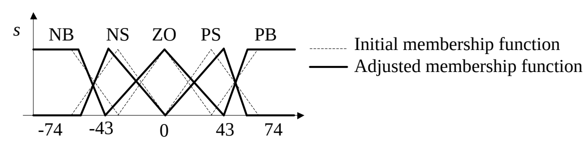

Fig. 15. The membership functions of the fuzzy variables before and after learning.

TABLE XVI
The fuzzy rule base of ΔK_P before and after learning

	ΔK_P				
s	NB	NS	ZO	PS	PB
u_{old}	PB	NS	ZO	PS	NB
u_{new}	PB	NB	PS	NB	NS

TABLE XVII
The fuzzy rule base of ΔK_I before and after learning

	ΔK_I				
s	NB	NS	ZO	PS	PB
u_{old}	NB	NS	ZO	PS	PB
u_{new}	NB	NB	PS	NB	PB

TABLE XVIII
The fuzzy rule base of ΔK_D before and after learning

	ΔK_D				
s	NB	NS	ZO	PS	PB
u_{old}	PS	NS	NS	ZO	PB
u_{new}	NB	NB	ZO	NB	PS

In the experiment of position control command response, there are mainly four comparative control methods: (1) PID, (2) fuzzy PID (FSPID), (3) fuzzy PID based on the membership functions of Q-learning (QMFSPID), and (4) Fuzzy PID based on Q-learning membership functions and fuzzy rule base (QMFFSPID). The position control command response is shown in Fig. 16, and the comparison of the Integral Square Error (ISE) and settling time is shown in Table XIX.

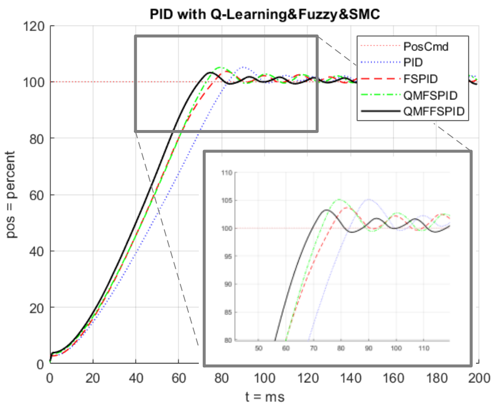

Fig. 16. Position control command response of voice coil motor.

TABLE XIX
THE COMPARISON BETWEEN INTEGRAL SQUARE ERROR (ISE) AND SETTLING TIME OF VOICE COIL MOTOR

	PID	FSPID	QMFSPID	QMFFSPID
ISE	341.22	308.77	305.64	**286.30**
Settling Time (ms)	98	88	87	**80**

B. Speed Control of the Brushless DC Motor Experimental Platform

In the process of learning and adjusting the membership functions of the fuzzy variable of the fuzzy controller by Q-learning, it is mainly to sequentially order 80% speed command, 0% speed command, -80% speed command and 0% speed command on the motor of the brushless DC motor experiment platform. There are 4 speed commands in total, and these 4 commands are repeated 100 times in sequence. The membership functions of the fuzzy variables before and after learning are shown in Fig. 17. The fuzzy rule base of ΔK_P, ΔK_I, and ΔK_D before and after learning are shown in Table XX, Table XXI, and Table XXII.

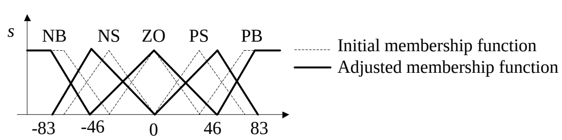

Fig. 17. The membership functions of the fuzzy variables before and after learning.

TABLE XX
THE FUZZY RULE BASE OF ΔK_P BEFORE AND AFTER LEARNING

ΔK_P					
s	NB	NS	ZO	PS	PB
u_{old}	PB	NS	ZO	PS	NB
u_{new}	PB	ZO	PB	ZO	NB

TABLE XXI
THE FUZZY RULE BASE OF ΔK_I BEFORE AND AFTER LEARNING

ΔK_I					
s	NB	NS	ZO	PS	PB
u_{old}	NB	NS	ZO	PS	PB
u_{new}	NS	PS	PS	PB	PB

TABLE XXII
THE FUZZY RULE BASE OF ΔK_D BEFORE AND AFTER LEARNING

ΔK_D					
s	NB	NS	ZO	PS	PB
u_{old}	PS	NS	NS	ZO	PB
u_{new}	NB	PS	NS	PB	NB

In the experiment of speed control command response, there are mainly 4 comparative control methods: (1) PID, (2) FSPID, (3) QMFSPID, and (4) QMFFSPID. The speed control command response is shown in Fig. 19, and the comparison between the ISE and the settling time is shown in Table XXIII.

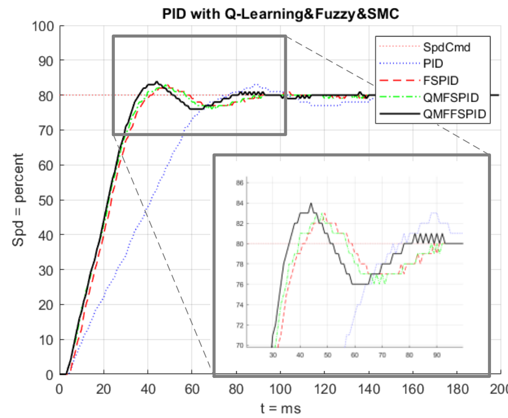

Fig. 19. Speed control command response of brushless DC motor.

TABLE XXIII
THE COMPARISON BETWEEN THE ISE AND THE SETTLING TIME OF BRUSHLESS DC MOTOR

	PID	FSPID	QMFSPID	QMFFSPID
ISE	154.41	98.23	89.79	**88.73**
Settling Time (ms)	92	55	50	**48**

C. Motor Speed Control of Two-Wheeled Self-Balancing Vehicle

In the process of learning and adjusting the membership functions of the fuzzy variable of the fuzzy controller by Q-learning, it is mainly to sequentially order 80% speed command, 0% speed command, -80% speed command and 0% speed command, a total of 4 speed commands. Then repeat these 4 commands 100 times in sequence. The membership functions of the fuzzy variables before and after learning are shown in Fig. 20. The fuzzy rule base of ΔK_P, ΔK_I, and ΔK_D before and after learning are shown in Table XXIV, Table XXV, and Table XXVI.

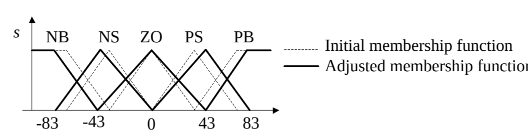

Fig. 20. The membership functions of the fuzzy variables before and after learning.

TABLE XXIV
THE FUZZY RULE BASE OF ΔK_P BEFORE AND AFTER LEARNING

ΔK_P					
s	NB	NS	ZO	PS	PB
u_{old}	PB	NS	ZO	PS	NB
u_{new}	PB	PS	PB	ZO	NS

TABLE XXV
THE FUZZY RULE BASE OF ΔK_I BEFORE AND AFTER LEARNING

ΔK_I					
s	NB	NS	ZO	PS	PB
u_{old}	NB	NS	ZO	PS	PB
u_{new}	NS	ZO	PS	PB	NB

TABLE XXVI
THE FUZZY RULE BASE OF ΔK_D BEFORE AND AFTER LEARNING

ΔK_D					
s	NB	NS	ZO	PS	PB
u_{old}	PS	NS	NS	ZO	PB
u_{new}	**NS**	NS	**PS**	**PB**	**NS**

In the experiment of speed control command response, there are mainly 4 comparative control methods: (1) PID, (2) FSPID, (3) QMFSPID, and (4) QMFFSPID. The speed control command response is shown in Fig. 21, and the comparison between the ISE and the settling time is shown in Table XXVII.

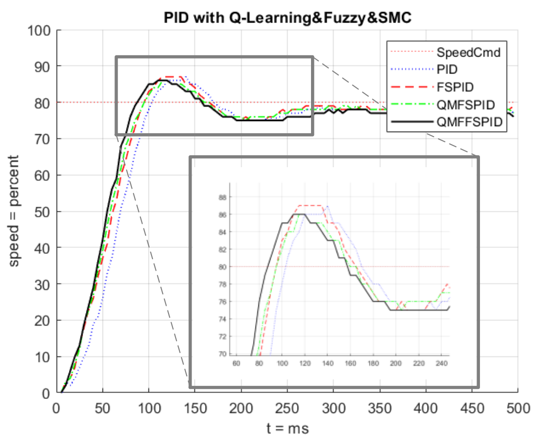

Fig. 21. Motor speed control command response of two-wheeled self-balancing vehicle.

TABLE XXVII
THE COMPARISON BETWEEN THE ISE AND THE SETTLING TIME OF TWO-WHEELED SELF-BALANCING VEHICLE

	PID	FSPID	QMFSPID	QMFFSPID
ISE	64.03	53.47	51.71	**50.03**
Settling Time (ms)	160	155	145	**140**

IV. CONCLUSIONS

In this paper, a framework with the Q learning algorithm is proposed to adjust the membership functions of fuzzy variables and the rule base of the fuzzy PID controller. Moreover, in order to improve the learning efficiency and reduce the complexity of the Q learning algorithm for learning the parameters of the fuzzy controller, the sliding mode concept is used to reduce the number of parameters required by the fuzzy controller. From the experimental results of the voice coil motor, brushed DC motor, and brushless DC motor, we can see that the proposed method does have better motor control results.

REFERENCES

[1] K. J. Astrom and T. Hagglund, *PID controllers: theory, design, and tuning*, Research Triangle Park, NC: Instrument society of America, 1995.
[2] Mamdani and H. Ebrahim "Application of fuzzy algorithms for control of simple dynamic plant," *Proceedings of the institution of electrical engineers*, vol. 121, no. 12, pp. 1585-1588, 1974.
[3] R. R. Yager and P. Dimitar, "Essentials of fuzzy modeling and control," *SIGART Bulletin*, vol. 6, no. 4, pp. 22, 1994.
[4] J. X. Xu, C. C. Hang, and C. Liu, "Parallel structure and tuning of a fuzzy PID controller," *Automatica*, vol. 36, no. 5, pp. 673-684, 2000.
[5] M. Mizumoto, "Realization of PID controls by fuzzy control methods," *Fuzzy sets and systems*, vol. 70, no. 2-3, pp. 171-182, 1995.
[6] S. J. Qin and G. Borders, "A multiregion fuzzy logic controller for nonlinear process control," *IEEE Transactions on Fuzzy Systems*, vol. 2, no. 1, pp. 74-81, 1994.
[7] J. X. Xu, Y. M. Pok, C. Liu, and C. C. Hang, "Tuning and analysis of a fuzzy PI controller based on gain and phase margins," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 28, no. 5, pp. 685-691, 1998.
[8] S. Z. He, S. Tan, F. L. Xu, and P. Z. Wang, "Fuzzy self-tuning of PID controllers," *Fuzzy sets and systems*, vol. 56, no. 1, pp. 37-46, 1993.
[9] Z. Y. Zhao, M. Tomizuka, and S. Isaka, "Fuzzy gain scheduling of PID controllers," *IEEE transactions on systems, man, and cybernetics*, vol. 23, no. 5, pp. 1392-1398, 1993.
[10] R. S. Sutton, A. G. Barto, and R. J. Williams, "Reinforcement learning is direct adaptive optimal control," *IEEE Control Systems Magazine*, vol. 12, no. 2, pp. 19-22, 1992.
[11] R. S. Sutton, A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.
[12] A. Bonarini, "Evolutionary learning of fuzzy rules: competition and cooperation," *Fuzzy Modelling*, vol. 7, pp. 265-283, 1996.
[13] T. Kawabe, T. Tagami, and T. Katayama, "A genetic algorithm based minimax optimal design of robust I-PD controller," *UKACC International Conference on Control. Control '96*, pp. 436-441, 1996.
[14] C. K. Chiang, H. Y. Chung, and J. J. Lin, "A self-learning fuzzy logic controller using genetic algorithms with reinforcements," *IEEE Transactions on Fuzzy Systems*, vol. 5, no. 3, pp. 460-467, 1997.
[15] E. S. Low, P. Ong, and K. C. Cheah, "Solving the optimal path planning of a mobile robot using improved Q-learning," *Robotics and Autonomous Systems*, vol. 115, pp. 143-161, 2019.
[16] V. T. Aghaei, A. Onat, I. Eksin, and M. Guzelkaya, "Fuzzy PID controller design using Q-learning algorithm with a manipulated reward function," *2015 European control conference (ECC), IEEE*, pp. 2502-2507, 2015.
[17] P. Kofinas and A.I. Dounis, "Online tuning of a PID controller with a fuzzy reinforcement learning MAS for flowrate control of a desalination unit," *Electronics*, vol. 8, no. 2, pp. 231, 2019.
[18] J. H. Li, Design of Fuzzy Sliding-Mode Controllers and Their Applications to a Class of Mechatronic Systems, National Cheng Kung University, 2003.
[19] S. W. Wu, Design of a Fuzzy Sliding-Mode Controller for Sensorless V/f Controlled Induction Motor, Chien Hsin University of Science and Technology, 2011.

CHING-CHANG WONG received the B.S. degree from the Department of Electronic Engineering, Tamkang University, Taiwan, in 1984, and the M.S. and Ph.D. degrees from the Department of Electrical Engineering, Tatung Institute of Technology, Taiwan, in 1986 and 1989, respectively. In 1989, he joined the Department of Electrical and Computer Engineering, Tamkang University (TKU), where he served as the Department Chairman from 2006 to 2010. In 2007, he established the Robotics Engineering Institute. He is currently a Distinguished Professor. In 2011, he established the Intelligent Automation and Robotics Center. His current research interests include intelligent control, humanoid robot, mobile robot manipulator, and deep reinforcement learning for robotic applications.

I-SHEN YEH was born in Taoyuan, Taiwan, in 1993. He received his B.S. and Ph.D. degrees in electrical and computer engineering from Tamkang University, Taipei, Taiwan, in 2015 and 2020, respectively. His major research interests include robot manipulator, robotic applications, and motor controller.

SHAO-YU CHIEN was born in Keelung, Taiwan, in 1995. He received the B.S. degree and the M.S. degrees from the Department of Electrical and Computer Engineering, Tamkang University, Taiwan, in 2017 and 2019, respectively, where he is currently pursuing the Ph.D. degree. His major research interests include robot manipulator, robotic applications, and machine learning.