

## Research Article

### Reinforcement Learning with FCMAC for TRMS Control

Jih-Gau Juang and Yi-Chong Chiang

Department of Communications, Navigation and Control Engineering, National Taiwan Ocean University, Keelung, 20224, Taiwan

**Abstract:** This study proposes an intelligent control scheme that integrate reinforcement learning in Fuzzy CMAC (FCMAC) for a Twin Rotor Multi-input and multi-output System (TRMS). In the control design, fuzzy CMAC controller is utilized to compensate for PID control signal and the reinforcement learning refines the compensation to the control signal. CMAC with fuzzy system has better performance than the conventional CMAC in TRMS attitude tracking control. With reinforcement learning, the proposed control scheme provides even better performance and control for the TRMS.

**Keywords:** Fuzzy CMAC, PID control, reinforcement learning, twin rotor MIMO system

#### INTRODUCTION

Reinforcement learning is used to the problem that faces an agent who wishes to learn its action selection function that can lead to an optimal reward, through trial-and-error interactions in a dynamic environment (Yeh, 2009). The learner is not told which actions to take, as in most forms of machine learning, but instead must discover which actions yield the most reward by trying them (Sutton and Barto, 1998). In the most interesting and challenging cases, actions may affect not only the immediate reward but also the next situation. Reinforcement learning is defined not by characterizing learning methods, but by characterizing a learning problem. The basic idea is simply to capture the most important aspects of the problem facing a learning agent interacting with its environment to achieve a goal. The agent must be able to sense the state of the environment to some extent and must be able to take actions that affect the state. To obtain better reward, a reinforcement learning agent must prefer actions that it has tried in the past and found to be effective in producing reward. Reinforcement learning has been successfully applied to many control problems. It is one of the best learning algorithms for systems without desired values (Anderson *et al.*, 2007; Valasek *et al.*, 2008).

In recent years, intelligent system research has become main stream in the field of control engineering. Control design uses the concept of artificial intelligence is often called intelligent control. Although the intelligent controller has been widely applied to many aspects, but the PID controller is still the most common controller in the industry. PID control structure is

simple and parameter adjustment method is easy (Kuo, 1995; Krohling and Rey, 2001). But the disadvantage is that suitable values for control parameters are not easy to obtain. Sometimes the optimization process is very time consuming. With the development of intelligent control, the shortcomings of the PID control can be overcome. There are many types of intelligent algorithms, such as neural networks, fuzzy logic systems and genetic algorithms (Juang *et al.*, 2011). In artificial neural networks, CMAC (Albus, 1975) has been used for many applications. Conventional neural networks and CMAC are different on weight updating rule. CMAC has fast convergence speed and low computation time. Hardware implementation of CMAC neural network is easy to achieve. In this study, a fuzzy reasoning mechanism to obtain the memory mapped values of CMAC is applied, which is called FCMAC. Adaptive learning rate of CMAC can be obtained (Yang and Juang, 2009) by the Lyapunov stability theory. This study presents several control methods to control a TRMS (Feedback Co, 1998). Controller design is based on reinforcement learning, CMAC, FCMAC and PID control. The controlled plant, TRMS, is a high order nonlinear system with cross coupling effect. The goal of this study is to make TRMS move fast, accurate and stable to the desired attitude.

#### METHODOLOGY

**System description:** Figure 1 shows a diagram of the laboratory set-up. There are four variables that are used to describe the state of the beam and two additional states variables are the angular velocity of the rotors, measured by tachometers coupled with the driving DC

**Corresponding Author:** Jih-Gau Juang, Department of Communications, Navigation and Control Engineering, National Taiwan Ocean University, Keelung, 20224, Taiwan

This work is licensed under a Creative Commons Attribution 4.0 International License (URL: <http://creativecommons.org/licenses/by/4.0/>).

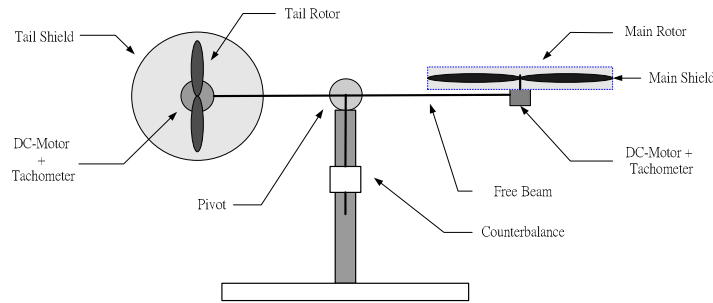


Fig. 1: The laboratory set-up TRMS

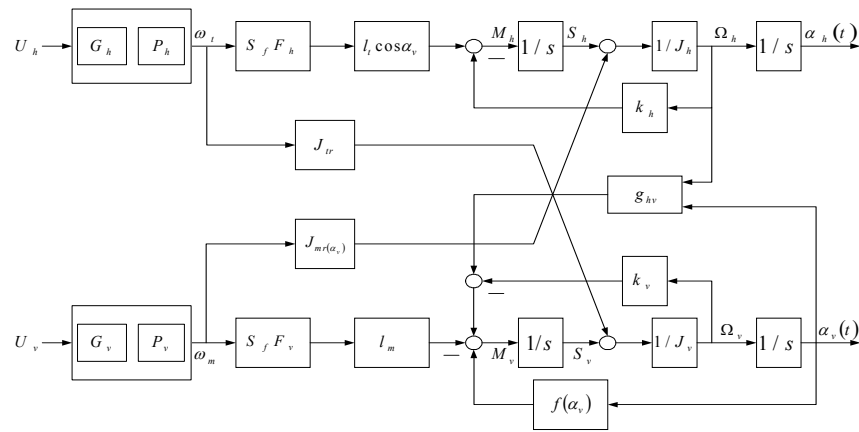


Fig. 2: Block diagram of TRMS model

motors. The horizontal and vertical angles and two corresponding angular velocities are measured by position sensors in the pivot. The two propellers are perpendicular to one another and they are built on a beam pivot that can rotate freely on the horizontal plan and vertical plane. By changing the input voltage on beam pivot we can control the propeller's rotational speed and make it achieve the desired goal. Figure 2 shows a block diagram of the TRMS model:

where,

- $\alpha_h$  : The horizontal position of the TRMS beam
- $\alpha_v$  : The vertical position of the TRMS beam
- $\Omega_h$  : The horizontal angular velocity of the TRMS beam
- $\Omega_v$  : The vertical angular velocity of the TRMS beam
- $U_h$  : The horizontal DC-motor voltage input
- $U_v$  : The vertical DC-motor voltage input
- $G_h$  : The linear transfer functions of the tail rotor DC-motor
- $G_v$  : The linear transfer function of the main DC-motor

- $h$  : The nonlinear part of the DC-motor with tail rotor
- $v$  : The nonlinear part of the DC-motor with main rotor
- $w_h$  : The rotational speed of the tail rotor
- $w_v$  : The rotational speed of the main rotor
- $F_h$  : The nonlinear function of the aerodynamic force from the tail rotor
- $F_v$  : The nonlinear function of the aerodynamic force from the main rotor
- $l_h$  : The effective arm of aerodynamic force from the tail rotor
- $l_v$  : The effective arm of aerodynamic force from the main rotor
- $J_h$  : The nonlinear function of the moment of inertia with respect to the horizontal axis
- $J_v$  : The nonlinear function of the moment of inertia with respect to the vertical axis
- $M_h$  : The horizontal tuning torque
- $M_v$  : The vertical tuning torque
- $K_h$  : The horizontal angular momentum

- $K_v$  : The vertical angular momentum
- $f_h$  : The moment of friction force in the horizontal plane
- $f_v$  : The moment of friction force in the vertical plan
- $J_{hv}$  : The horizontal angular momentum from the main rotor
- $J_{vh}$  : The vertical angular momentum from the tail rotor
- $f$  : The vertical tuning moment from counterbalance
- $g_{hv}, g_{vh}$  : The nonlinear function of the reaction tuning moment

**Controller design:** Figure 3 shows the proposed reinforcement learning with FCMAC control scheme. The reinforcement learning is applied to compute a compensation for FCMAC command. The reinforcement learning supervises the FCMAC controller by error between system output and reference and then gives rewards to FCMAC command. This study presents two ways of obtaining the rewards.

**Addition and subtraction:**

- If  $e > t_+$ , then  $g(k+1) = g(k) - \rho$
- If  $e < t_-$ , then  $g(k+1) = g(k) + \rho$

where,

- $e$  : System error
- $t_+$  &  $t_-$  : Thresholds of error
- $u_c$  : CMAC control command
- $g$  : A gain to the  $u_c$
- $\rho$  : The reward to  $g$
- $k$  : The time index

The output of CMAC is then multiplied by  $g(k+1)$  and new  $u_c$  can be obtained as follow:

$$u_c = g(k+1)u_c$$

**Use genetic algorithm to search reward:** We use a Real-valued Genetic Algorithm (RGA) to search parameters of reinforcement learning. Initial settings of genetic algorithm are: number of patterns is 10, number of generations is 50, mating rate is 0.8 and mutation rate is 0.2. Fitness function is given as follow:

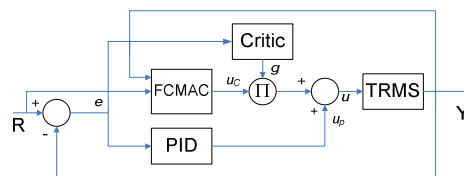


Fig. 3: Reinforcement FCMAC control scheme

$$fit(g) = \frac{1}{1 - \left( \frac{u_c}{u_c g - e} \right)}$$

In *A*, to set an error threshold, if the system error is greater than a preset threshold, the control gain will be punished by  $\rho$ . On the contrary, if the error is less than the threshold, the gain will be rewarded by  $\rho$ . This method is based on the relationship between the reference and system output. If control command is too small, then the control command will have the addition operation and vice versa, if control command is too large, then the control command will have the subtraction operation. In *B*, a RGA is used to search appropriate gain value. The most important thing is how to determine the fitness function. We choose RGA with appropriate initial setting: pattern number, generation number, mating, mutation, then the genetic algorithm can replace trail-and-error used in *A*, to obtain optimal control gain.

Cerebellar model articulation controller (Glanz *et al.*, 1991) is a complex nonlinear expression of the form of lookup-table and can be considered as a kind of neural network. This network may change form content through the learning algorithm. CMAC makes system's input state as an indicator and stores related information in a group of memory cell. It is essentially a kind of mapping the complex nonlinear function into a look-up table. Specifically, the input space is divided into many sub-blocks. Each block specifies a real memory location. Each block is used to learn the information distributed to the adjacent block's storage position. CMAC has been recognized as class associative memory neural network. It can learn any multi-dimensional nonlinear mapping. CMAC algorithm can be effectively used for nonlinear function approximation, dynamic modeling, control system design, etc. Compared with other neural networks, CMAC is based on local learning. Its information is stored in partial structure and only few weights are updated at each iteration and this makes it appropriate for real-time control. As a nonlinear approximation, it is insensitive to the order of learning data. The input maps

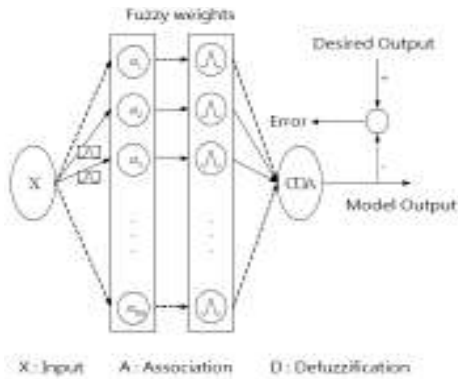


Fig. 4: The conceptual diagram of FCMAC

to the actual memory of the  $N$  units. Each stored unit with corresponding weight value, CMAC output of the  $N$  units is a weighted sum of the actual memory cell. The difference between FCMAC and CMAC is in the mapping part, as shown in Fig. 4. The mapping of FCMAC uses fuzzy reasoning mechanism to get mapping value. The mapping value can help CMAC controller to reduce computing times. The fuzzy reasoning mechanism does not increase CMAC updating time and computing time. On the contrary, it reduces updating time and computing time.

FCMAC includes two kinds of operations that calculate the output result and adjust the weights. The mapping  $A \rightarrow Y$  is to compute the output as:

$$y(x) = \mathbf{a}^T \mathbf{w}(x) = \sum_j^N a_j(x) w_j, j = 1, 2, \dots, N$$

where,

- $w_j$  : The weight of the  $j$ -th storage hypercube
- $a_j(x)$  : A factor obtained from fuzzy reasoning mechanism indicating

whether the  $j$ -th storage hypercube is addressed by the input  $x$ . It is to modify the weight of storage hypercube according to the error between the desired output and the real output, in the stage of network learning in FCMAC. Its weight updating rule is:

$$w_j^{(i)} = w_j^{(i-1)} + \Delta w_j^{(i-1)} = w_j^{(i-1)} + \frac{\alpha}{m} a_j (y_d - \sum_{j=1}^N w_j^{(i-1)} a_j)$$

where,

- $i$  : The  $i$ -th iteration
- $y_d$  : The desired output
- $m$  : The number of addressed hypercube
- $\alpha$  : The learning rate

In order to improve the shortcoming of conventional CMAC on the crisp relation, adaptive

learning rate is introduced. The adaptive CMAC uses a set of adaptive learning rates that are derived from the Lyapunov stability theory:

$$\frac{2m}{\|\mathbf{a}\|^2} > \alpha > 0$$

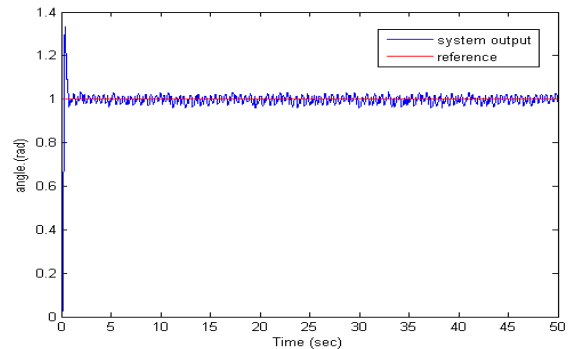
where,  $\mathbf{a}$  is the vector of  $a_j(x)$ .

The conventional CMAC has good performance in decoupled system of the TRMS, but the control of cross-coupling system can not be achieved. Therefore this study presents a FCMAC and PID control scheme to achieve cross-coupling control. The FCMAC realizes the feed-forward control, which is used to compensate for the control signal that is applied to the TRMS. The PID controller realizes the feedback control that guarantees system's stability and disturbance resistance.

### EXPERIMENTAL RESULTS

In cross-coupled condition, the task of control is to overcome the cross-coupling influence between vertical and horizontal axes. Figure 5 to 7 show the step, sine wave and square wave responses in cross-coupled condition using CMAC controller, respectively.

2-DOF horizontal simulation



2-DOF vertical simulation

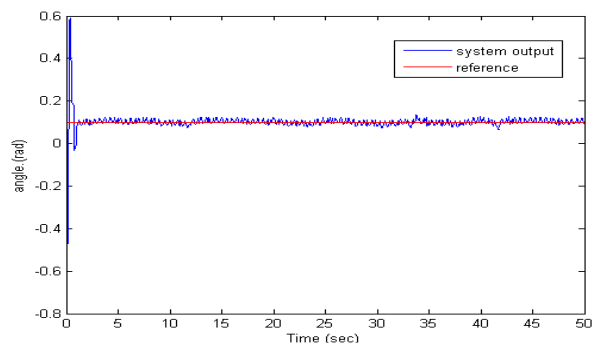


Fig. 5: Step response using CMAC

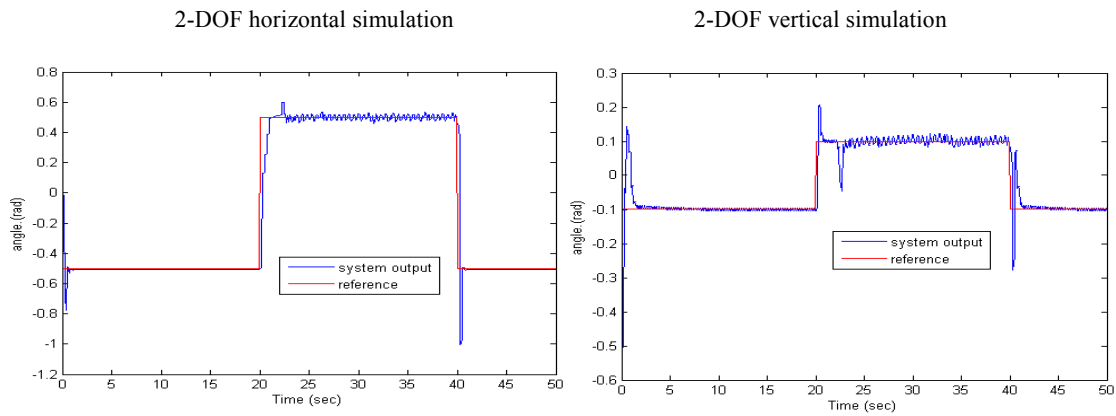


Fig. 6: Square wave response using CMAC

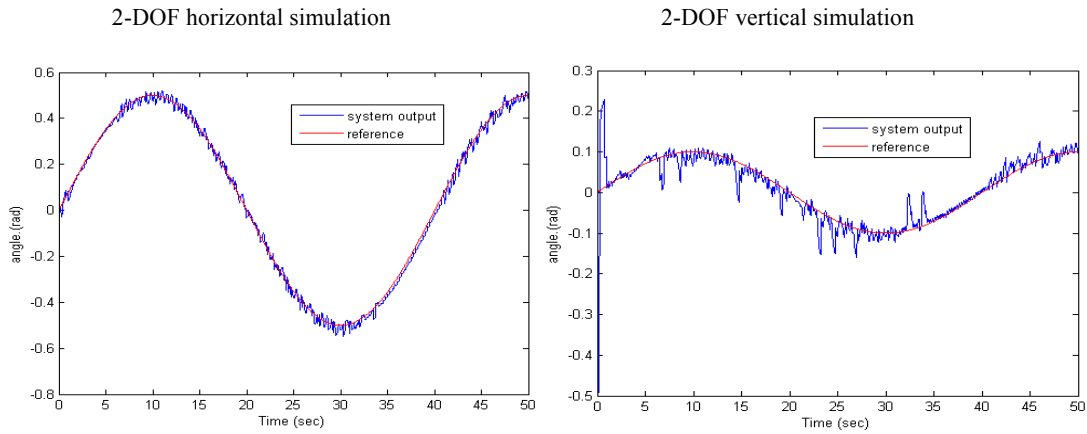


Fig. 7: Sine wave response using CMAC

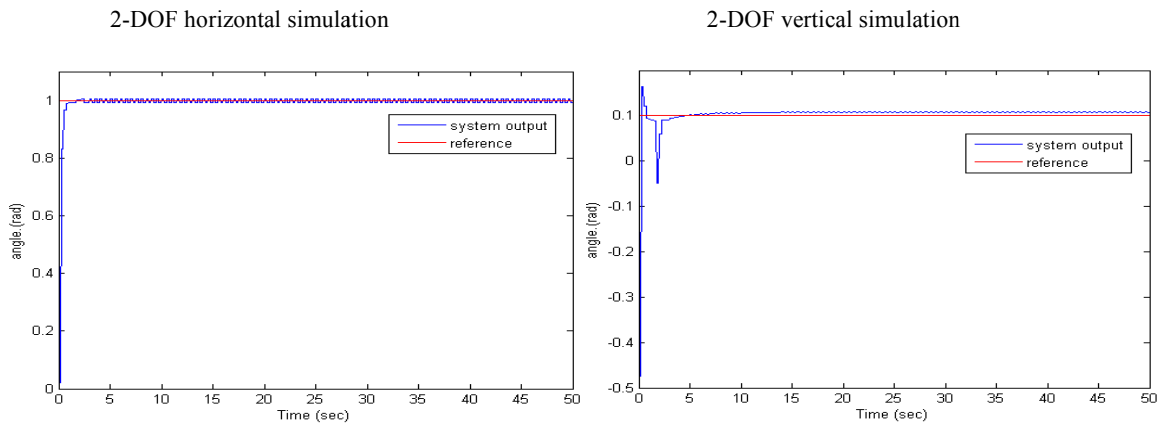


Fig. 8: Step response using FCMAC

Figure 8 to 10 show the step, sine wave and square wave responses in cross-coupled condition using FCMAC controller. Figure 11 to 13 show the step, sine wave and square wave responses in cross-coupled condition using

reinforcement FCMAC controller. Compare control performances of CMAC, FCMAC (Chiang and Juang, 2011) and reinforcement FCMAC in the cross-coupled condition by total tracking errors in 50 sec, as shown in

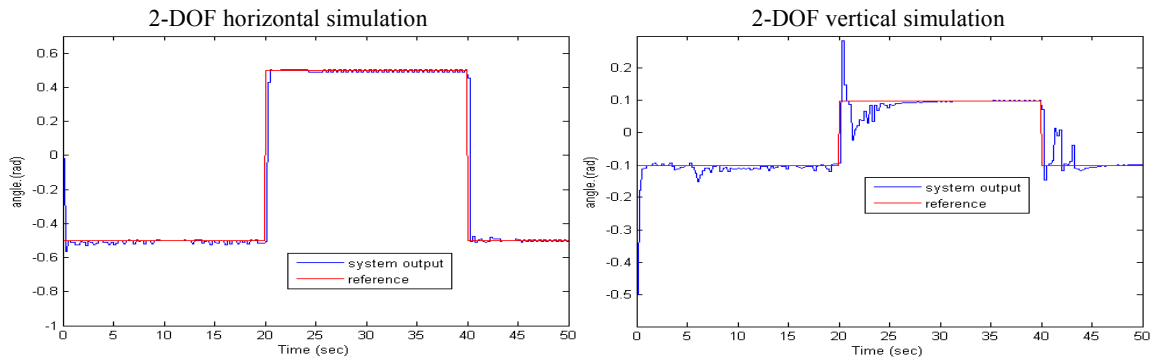


Fig. 9: Square wave response using FCMAC

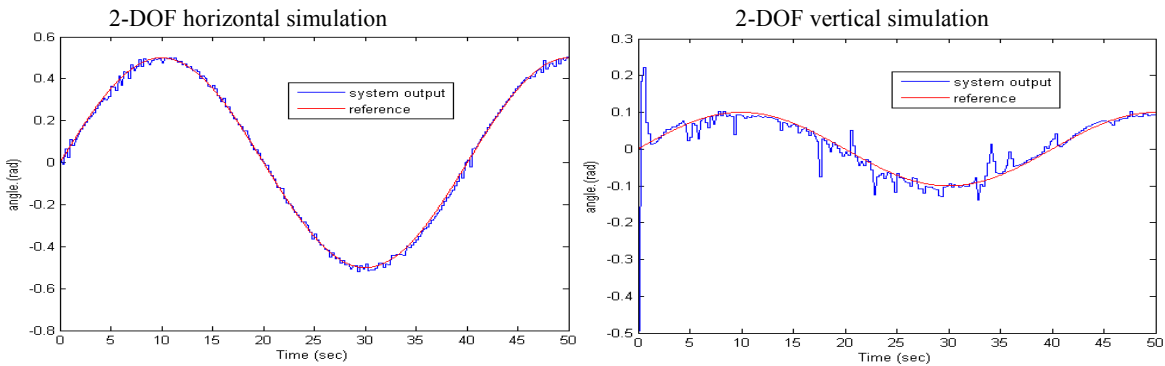


Fig. 10: Sine wave response using FCMAC

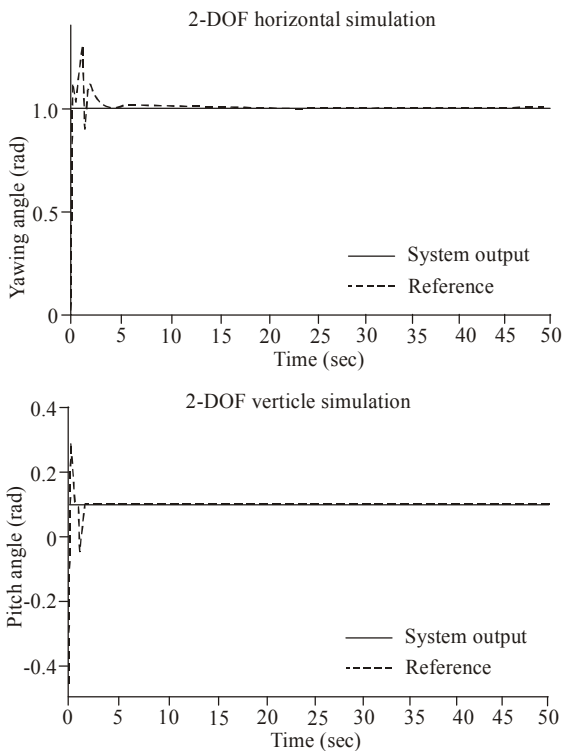


Fig. 11: Step response using reinforcement FCMAC

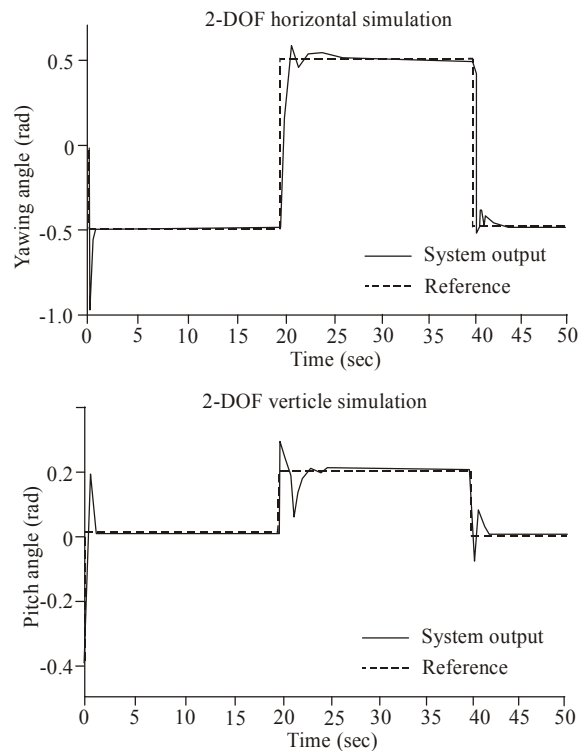


Fig. 12: Square wave response using reinforcement FCMAC

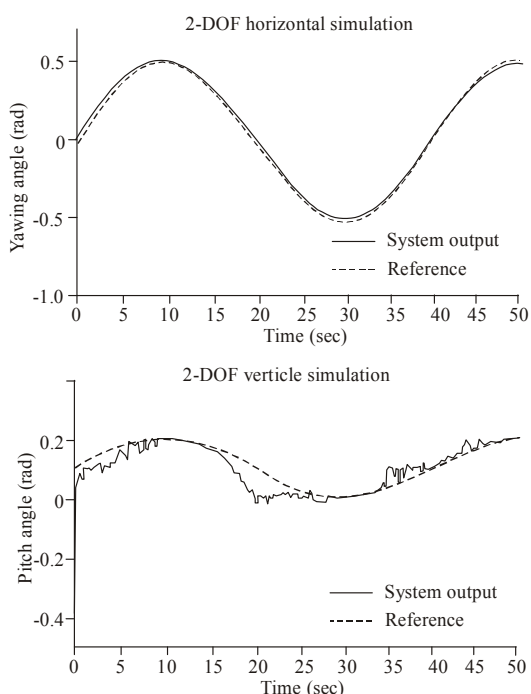


Fig. 13: Sine wave response using reinforcement FCMAC

Table 1: Comparison of CMAC, FCMAC, and reinforce FCMAC

Type		Step response total error	Square wave total error	Sine wave total error
CMAC	Horizontal	21.94	29.45	16.44
	Vertical	14.66	15.08	18.19
FCMAC	Horizontal	21.87	28.90	18.09
	Vertical	12.91	15.35	20.77
Reinforce FCMAC	Horizontal	18.66	30.13	13.73
	Vertical	7.410	11.38	23.27

Table 1, the proposed reinforcement FCMAC control scheme has better average performance than other controllers.

### CONCLUSION

This study presents the use of conventional CMAC, fuzzy CMAC and reinforcement fuzzy CMAC controllers to control a twin rotor MIMO system. Simulation results show that the proposed reinforcement FCMAC controller has better performance in system stability and attitude tracking control. In reinforcement learning control, compensation gain obtained by the addition and subtraction method provides worse compensation to the system, even worse than the CMAC control. By the use of genetic algorithm to search the

reward the control system performance is improved dramatically.

### REFERENCES

Albus, J.S., 1975. A new approach to manipulator control: The Cerebellar Model Articulation Controller (CMAC). *J. Dyn. Syst. Meas. Control*, Trans. ASME, 97: 220-227.

Anderson, C.W., P.M. Young, M.R. Buehner, J.N. Knight, K.A. Bush and D.C. Hittle, 2007. Robust reinforcement learning control using integral quadratic constraints for recurrent neural networks. *IEEE T. Neural Networ.*, 18(4): 993-1002.

Chiang, Y.C. and J.G. Juang, 2011. Application of fuzzy CMAC to twin rotor MIMO system. *Proceeding of National Symposium on System Science and Engineering*, ID-193.pp:

Feedback Co, 1998. *Twin Rotor MIMO System 33-220 User Manual*, 3-000M5.

Glanz, F.H., W.T. Miller and L.G. Kraft, 1991. An overview of the CMAC neural network. *IEEE Conference on Neural Networks for Ocean Engineering*, pp: 301-308.

Juang, J.G., W.K. Liu and R.W. Lin, 2011. A hybrid intelligent controller for a twin rotor MIMO system and its hardware implementation. *ISA T.*, 50(4): 609-619.

Krohling, R.A. and J.P. Rey, 2001. Design of optimal disturbance rejection PID controllers using genetic algorithms. *IEEE T.*

Kuo, B.C., 1995. *Automatic Control Systems*. 7th Edn., Prentice-Hall, Englewood Cliffs, NJ.

Sutton, R.S. and A.G. Barto, 1998. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge.

Valasek, J., J. Doebbler, M.D. Tandale and A.J. Meade, 2008. Improved adaptive-reinforcement learning control for morphing unmanned air vehicles. *IEEE T. Syst. Man Cy. B.*, 38(4): 1014-1020. *Evolut.*, 5(1): 78-82.

Yang, T.C. and J.G. Juang, 2009. Aircraft landing control based on adaptive CMAC. *Proceeding of IEEE International Conference on Advanced Intelligent Mechatronics*, pp: 791-796.

Yeh, N.K., 2009. Dealing with perceptual aliasing by using pruning suffix tree memory in reinforcement learning. M.A. Thesis, National Tsing Hua University, Taiwan.