2013 Joint International Conference on Rural Information &
Communication Technology and Electric-Vehicle Technology (rICT & ICeV-T)
November 26-28, 2013, Bandung-Bali, Indonesia

Application of Reinforcement Learning on Self-Tuning PID Controller for Soccer Robot Multi-Agent System

Aulia el hakim[#1], Hilwadi Hindersah[#2], Estiko Rijanto[*3]
[#]*School of Electrical Engineering and Informatics, Bandung Institute of Technology,*
Ganesha Street 10, Bandung 40132, Indonesia,
[1]a_im07@yahoo.co.id, [2]hilwadi@stei.itb.ac.id
[*]*Research Center For Electrical Power and Mechatronics- LIPI,*
Jl.Sangkuriang, Bandung 40135, Indonesia,
[3]estiko.rijanto@lipi.go.id

Abstract**—** In the soccer robot game, quickly and accurately move is a very necessary. Proportional-Integral-Derivative (PID) controllers are applied to overcome it. However, to obtain the optimal control of robots, required tuning PID of parameters (Kp, Ki, and Kd) that not easy. In this research used reinforcement learning algorithm with Q-Learning method to determine the parameters in the process of self-tuning PID control.

To quantify the results that obtained, it would require a comparison between the use of RL algorithms on self-tuning PID with PID conventional usage by tuning using Ziegler-Nichols oscillation method, and compared to controls that have been implemented in the YSR-A Yujin Robotics. Test was conducted by moving the 5 robot with combination 3 robots of red and 2 white robot toward a particular point and to play with a certain angle. From the test results is known that the values obtained using the RL algorithm parameters PID control output which produces the most stable and have 1.875 faster response than using the methods applied YSR-A Yujin Robotis and have 1.5 faster response compared by using the Ziegler-Nichols oscillation method.

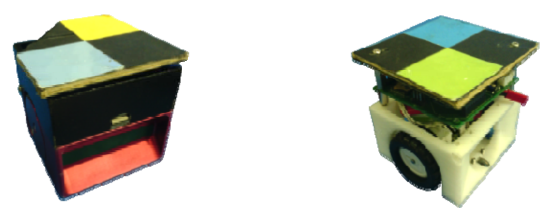
Keywords : Soccer robot,PID, Reinforcement learning, Q-Learning, self-tuning PID

I. INTRODUCTION

An efforts to promote science and technology in the field of artificial intelligence and robotics to the community, especially the younger generation has been developed and carried out. One form of such recognition effort is to use soccer robot games. The idea of using football to promote science and technology of artificial intelligence and robotics have presented starting in the 90s [1]. Robot soccer is an interdisciplinary research environment that combines perception, dynamic decision-making, communication between the robot system and sub-system control of movement [1] [2]. And the latter is one of the most important parts of robot soccer [2]. Sub control system aims to regulate the movement and the robot rotates quickly and accurately [3].

At this time, the conventional of PID control has been widely used both in industry and in the field of robot soccer control [4]. However application of conventional PID usually using fixed parameter, so its difficult to get the optimal controller effect[5]. There are 3 fixed parameter that must be sure in PID controller, include: K_p,K_i, dan K_d, i, It aims to obtain the optimal system response [6]. There are several methods developed in the process of tuning the parameters of PID include: ZieglerNichols method [7], [8], Cohen-Coon method [7], IAE and Itae [9], and the IMC (Internal Model Control) [7]. Generally, tuning methods were developed to control specific processes, so that the controller produces good operating only on the specific environment [10].

In the robot soccer mirosot divisions there are 5 robots in a team. And PID control is used so that each robot can move towards a certain position at a certain angle and turn quickly and accurately. In this research, a combination of two different types of robots in the team, which is red and white robot that can be seen in Figure 1. (a) and 1. (B).



(a) *(b)*

Fig.1 (a) YSR-A Robotics products, (b) ITB Robot Soccer team products

2 The robot system that has different specifications that can be seen in Table I. Under these conditions, the use of conventional PID will be hard to do, because there are 2 different types of robots in a team. Therefore reinforcement learning algorithms used to the process of self-tuning PID control with reference to the robot and the different conditions.

To measure the level of control that has been carried out optimization tests, the results obtained from testing the PID control will be compared with the results of control tests on YSR-Robotics and Ziegler-Nichols method to two.

II. LITERATURE

A. Soccer Robot

Soccer robot is a robot which combines artificial intelligence with mobile robotics technology which implemented in the soccer sports [12]. In essence, this research studied how mobile robots can be built and trained to be able to play the game of football [12]. The purpose of this research robot soccer is creating a robot soccer team that can implement strategies to the game of football, how to build artificial intelligence each agent that can think and work with other agents by consider the condition of the environment and how can the robot can beat opponents during the game [13].

B. Digital PID Controller

PID (Proportional Integral Derivative) controller is a control that is used to determine the precision of an instrumentation system by the characteristics of the feedback on the system. PID control is generally in Time continues domains, and the use of analog variables. Now digital technology has been used on a large scale and small as well as replace the analog control system into a digital control system. [17] There is easy to implement PID control in the digital version. This means that the control is in the domain discret PID. Future trends of digital control will be more widely used than by analog control due to the low cost [18]. To be able to understand the digital PID controller, as shown in Figure 2. And PID equation in the discrete domain [19] can be defined in Equation 1.

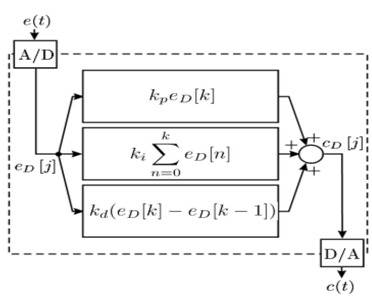


Fig. 2 Digital PID Controller [19]

$$c[k] = K\left(e[k] + \frac{T_c}{T_i}\sum_{n=0}^{k} e[n] + T_d\frac{e[k]-e[k-1]}{T_c}\right) \quad (1)$$

C. Tuning of PID control using the method of Ziegler - Nichols' Second

Tuning will be conducted in a closed loop which the reference input is a step function. This method of controlling the proportional controller only. Kp, is increased from 0 to a critical value of Kp, in order to obtain a continuous output oscillating by the same amplitude. The critical value is referred to as Ultimated Kp gain. Ziegler-Nichols proposes to determine the value of the parameters Kp, Ti and Td by meet the formula in table I. [9]

TABLE I
TUNING RULES OF ZIEGLER-NICHOLS SECOND METHOD

	P	PI	PID
Kp	$K_u / 2$	$2K_u / 5$	$3K_u / 5$
Ti	-	$4T_u / 5$	$T_u / 2$
Td	-	-	$3T_u / 25$

D. Reinforcement Learning

Reinforcement Learning is a process of learning what to do and how to map situations (mapping) of the existing situation in the environment (states) to the form of the action (behavior) in order to maximize reward [18]. In this study, which is used as agents of each robot soccer. Agent acting as a learner does not need to be notified whether the action should conducted, in other words, self-learning robot conducted learning from his experiences. When agents do the right thing based on the specified rule, the agent will get a reward, and vice versa. RL is generally composed of four basic components, namely Policy, Reward function, Value function, and the model of environment [19] [20].

E. Q-Learning

Q-Learning is an incremental dynamic programming procedure that determines the optimal policy in a step-by-step method [20].

$$\pi^*(s) = argmax_a\ Q(s_t, a_t)$$

$$Q(S_t, a_t) \leftarrow Q(S_t, a_t) + \alpha[r_{t+1} + \gamma\ max_a Q(S_{t+1}, a_{t+1}) - Q(S_t, a_t)] \quad (2)$$

these functions represent the reward value of taking action a result of agents of the state S which resulted in the displacement matters $S_{(t+1)}$. After getting the optimal Q-function, there are considerations optimization π^* (s) which is the maximum value of a state. π^* (s) is used to select the action in every state that passed.

F. Velocity equation of YSR-A robotics Yujin

At YSR-A yujin robotics is used only Proportional control so that the formulation of the linear velocity and angular robot as follows [11] [13]:

$$V_L = K_d.d_e - K_a.\theta_e$$

$$V_R = K_d.d_e + K_a.\theta_e \quad (3)$$

G. Velocity equation on research

This study used the PID control robot movements and refers to the formula (1), then the PID control obtained in the linear velocity and angular velocity as follows.

$$V = \left(Kpv.de + Kiv(de + de^{-1}).Ts + \left(\frac{Kdv}{Ts}\right) * (de - de^{-1})\right) \quad (4)$$

$$\omega = Kpa.\emptyset e + Kia(\emptyset e + \emptyset e^{-1}).Ts + \left(\frac{Kda}{Ts}\right)(\emptyset e - \emptyset e^{-1}) \quad (5)$$

With reference to the equation (4) and (5) then obtained in determining the PID Vr and Vl equation can be seen in the following formula.

$$Vl = \left(Kpd.de + Kid(de + de^{-1}).Ts + \left(\frac{Kdd}{Ts}\right)(de - de^{-1})\right) - \frac{L}{2}(Kpa.\emptyset e + Kia(\emptyset e + \emptyset e^{-1}).Ts + \left(\frac{Kda}{Ts}\right)(\emptyset e - \emptyset e^{-1}) \quad (6)$$

$$Vr = \left(Kpd.de + Kid(de + de^{-1}).Ts + (Kdd/Ts)(de - de^{-1})\right) + \frac{L}{2}(Kpa.\emptyset e + Kia(\emptyset e + \emptyset e^{-1}).Ts + (\frac{Kda}{Ts})(\emptyset e - \emptyset e^{-1}) \quad (7)$$

III. ANALYSIS AND DESIGN SYSTEM

A. Requirements analysis

In the robot soccer game, especially in the division soccer tournament Micro robot, the robot is required to perform the movement toward a particular point quickly and accurately, therefore we need a controller that can overcome these problems. In the control system of the robot soccer YSR-A control system that uses only include reinforcement parameter P, which can be divided into two parts, namely Kd parameters during the robot walks straight and Ka parameters when turning at a certain angle. Usage only one parameter in the PID control, can not make the optimal control of a the robot, so we need another parameter, namely parameters of I and D. Another issue that should be required solution is how to make the tuning of the PID gain parameters can be carried out correctly and produces the optimal value.

B. Design System

In the design of the control system is applied to the robot soccer can be divided into two parts, that is the design of reinforcement learning algorithm Q-Learning method and PID control design has been done with combination of reinforcement learning algorithms.

1) The design of reinforcement learning algorithms

In designing the reinforcement learning algorithm Q-lernning needs to be given first to a constant value of $\alpha = 0.5$ and $\gamma = 0.9$, the value is chosen at random with the requirements of a value between 0 to 1. The next step need to know state that will be passed by the learner. State on the study area based on the time taken by the robot and the the robot between the starting point of the robot is moving towards the target point. Division of the state can be seen in the table II.

In the process of learning the robot, the distance of covered robot to the target is the distance point error (de) received from the previous process, de can be divided into 3 areas: The first area with the value de> 35, the second area in the value of de > 8 and the third area with values de <= 8. Which can be seen in Figure 3.

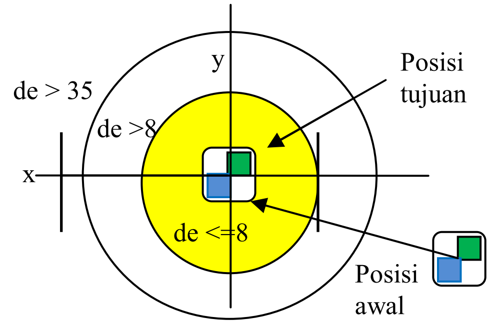


Fig. 3 Area division in the learning process

At the time of learning, the robot initial position is placed on de> 35, this is done so that the robot can be designed into any state. And can update the value of each action is executed.

TABLE II.
SITUATION ON STATE

No	Situasi	state	reward
1	Robot in the first area and the time taken <= 2 seconds	S1	2
2	Robot in the first area, 2 <time taken <4 seconds	S2	1
3	Robot in the first area and the time taken >= 4 seconds	S3	-10
4	Robot in the second area, the time taken <= 2 seconds	S4	4
5	Robot in the second area, 2 < the time taken <4 seconds	S5	2
6	Robot in the second area, the time taken> = 4 seconds	S6	-5
7	Robot in the third area, the time taken <= 2 seconds	S7	10
8	Robot in the third area, 2 < the time taken <4 seconds	S8	6
9	Robot in the third area, the time taken> = 4 seconds	S9	1

In every state that passed in reinforcement learning, there are actions that will be undertaken by the agents, the action that is the values of PID control parameters. In this study, the value of which may be divided into two parts, among others:

a) The action in PID parameters linear speed (v) :

- Kpd = { 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 }
- Kid = { 0.001, 0.002, 0.003, 0.004, 0.005, 0.006, 0.007, 0.008,000.9 }
- Kdd = { 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 }

b) The action in PID parameter angle speed (ω):

- Kpa = {0.01,0.02,0.03,0.04,0.05,0.06,.0.07,0.08,0.09 }
- Kia ={0.001,0.002,0.003,0.004,0.005,0.006,.0.007, 0.008, 0.009}
- Kda = { 0.001,0.002,0.003,0.004,0.005,0.006,0.007, 0.008,0.009}

The design of state control of the position of initial state that is state 1 to state 9 can be seen in Figure 4.

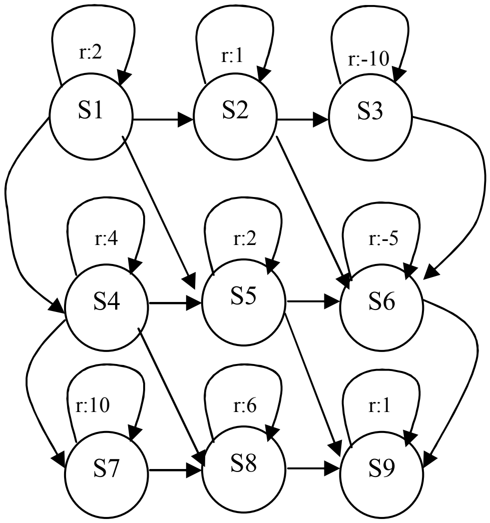


Fig. 4 design of state control of the robot soccer

2) Self-tuning PID control design

This research will use reinforcement learning algorithm Q-Learning methods in the process of self-tuning PID performed by each agents during moving to a certain position or a certain turn with swivel angle. Broadly speaking, the design of the control system can be seen in Figure 5

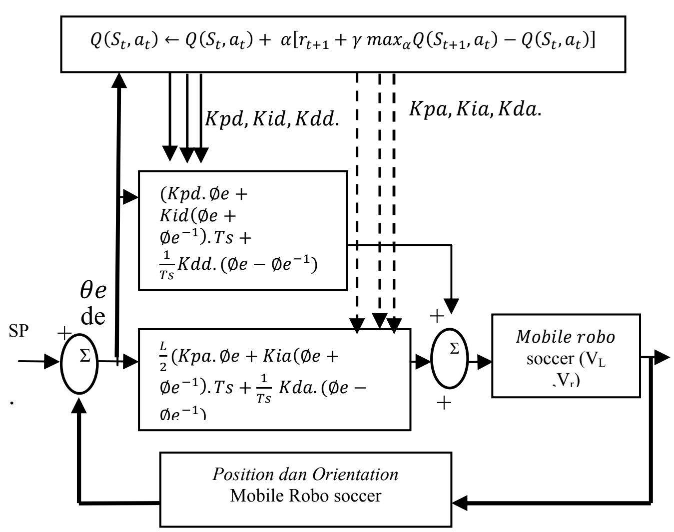


Fig.5 Robot soccer control system design

IV. IMPLEMENTATION AND TESTING

A. Implementation of reinforcement learning algorithm

One aspect that is needed in applying this algorithm is time and position information of each robot while in the field. Time data obtained from the data packets sent by the server. Then it used to a parameter in determining the state in accordance with the previous system design. After a known the value of time obtained from the packet data it can conducted implementation of the algorithm into the program. Broadly speaking, a program created in impementasi RL algorithms can be divided into two step, that is conditional state programs and programs to update the value of the Q-function of each state and action.

1) *Conditional state program*

On the conditioning state program, in addition to using time as a condition requirements, also required the value from the position of robot while in the field. It used to determining the value of de and teta e of the robot. Then from the de value will be used to determine the location of the area where the robot on the field in accordance with description in the system design. Areas where the robot along with the time that has been previously obtained will be used as a condition requirements by conducting operator "&&". After known to exist in which state, the robot find the value of Q (St, at) the greatest among the actions that are in the state. Having obtained the maximum action, then the action is included in the formulation of PID speed left and right wheels, and robot looking for the maximum value of Q (S_{t+1}, a).

2) *Q-Function update program*

After doing the conditional state program, the next step is make a listing program which are function to updating the value of Q-function that refer to the equation (2). In this programs, the value of Q-function obtained from the previous state will be changed with new value from Q(St, a) choiced in conditional state program.

B. Implementation of PID controller

With knowing the value of PID parameter (Kpd,Kid,Kdd, Kpa,Kia,Kda) from learning using RL algoithm , then the next step is applying those value to the control of velocity rigt (vr) and velocity left (vr) motor, refer to the equation (6) and (7) in previous chapter.

C. Testing

Testing was conducted based on analyzing the needs and design a system that has been described. The parts to be tested consists of: testing the proportional control of YS-A at 5 Robotics robot that have different characteristics with reference to the red robot and robot parameter white, Tests Against PID control with Ziegler-Nichols method for two robot on a combination of red and white. The final testing of the system of control that has been conducted using reinforcement learning. Scenario testing conducted by entering the command ToPosition (int withrobot, x, y) for each robot into the system's main function is the function My_Strategy (). ToPosition is a function in which there is code program for move ordered robot to the point x, and y destination.

1) Testing in in YSR-A Robotics control

Testing was conducted using 5 units robot with control parameter value is derived from the value of trial and error on a red robot with ID. 05. The parameter values for Kd = 0.8 and Ka = 0.3. on this test, five robot run together towards a certain point. Graph the results of testing against YSR-Robotics control can be seen in Fig. 6.

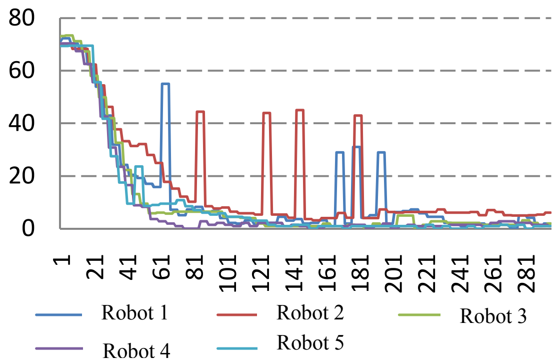


Fig. 6 Graphs from P-controller in 5 robots

From the test chart obtained by entering the parameter values obtained from ID.05, it is known that the White robot with ID.01 and ID.02 less stable, it is because the white robot characteristics and different red robot, so that the effect of control on The second robot to be different. And because the only form of control P control, the robot can not fix the steady state.

2) Testing control using Ziegler-Nichols second method.

Testing was conducted using 5 units robot with control parameter value is derived from the value of the tuning of a red robot with ID. 05. Parameter values obtained from the calculations in accordance with Table I. So we get the value: Kpd = 0,48, Kid = 0,367, Kdd = 0,1512, Kpa = 0,0222, Kia = 0,0273, Kda = 0,0433.
With the parameters obtained, it can be tested and the obtained results can be seen in Fig. 7.

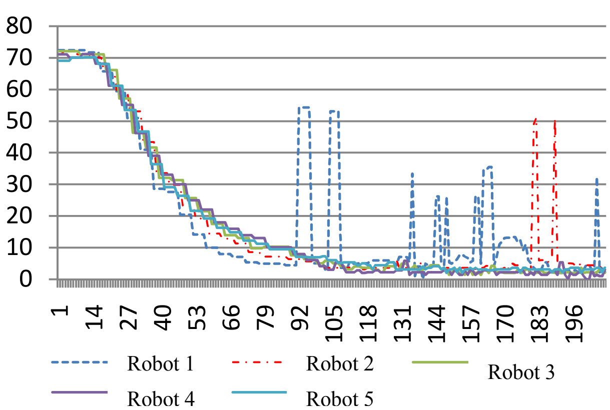


Fig. 7 Graphs from PID control by using the second Ziegler-Nichols method

On testing using the method of Ziegler nichols, have almost the same output data using YSR-A robotics control, but for robot with id: 3.4, 5 is more stable.

3) Testing parameters control using parameter that results from RL algorithm.

After doing a learning on red robot with ID.05, the PID parameter values obtained as follows:: Kpd= 0.4, Kid = 0.002, Kdd= 0.1, Kpa= 0,02, Kda= 0.002, Kia = 0.008. The parameters value are then inserted into the formulation of velocity (vl and vr) in ToPosition function.

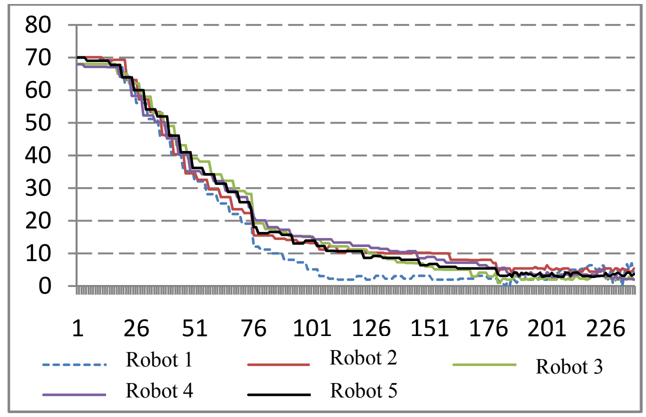


Fig. 8 Graphs from Learning PID control results

From graphic Fig. 8 can be analyzed that by using the parameters outcomes using RL, obtained a more stable control charts for all robots. It can be seen from the absence of large oscillations in each robot that run towards the expected position. And of drawing graphs 8 also found that all robots are already in the target position after a time of 181 units of time. After compared with the time taken by all the robots toward goal positions obtained from the graph of YSR-A Yujin robotics control and graphics using Ziegler-nichols tuning method, so travel time robot towards position obtained using control results from tuning using RL faster 15 time units or 1,875 seconds compared to the control using the YSR-A robotics, as well as a faster 12 time units or 1.5 seconds compared to the results using the control tuning using Ziegler-nichols.

V. CONCLUSION

After testing and analysis in the previous chapter, some conclusions can be drawn as follows.

1. PID control system developed in this study is a control system that can perform tuning The parameters reinforcement by itself. System will perform learning to obtain the optimal The parameters values.
2. The results of the design and implementation of PID The parameters tuning using reinforcement learning algorithm Q-learning has to be done properly.
3. Required learning by a relatively long period of time to obtain parameter values Kpd, Kid, Kdd, Kpa, Kia, and Kda that optimal, so if robots continue to run while the learning process is done, the parameters are sought to be more optimal.
4. From the graft position and angle control can be seen that the robot by using The parameters results learning can move to the target position with a more stable and 1.875 seconds faster than the control YSR-A yujin

Robotics and 1.5 seconds compared to the control results using the Ziegler-tuning nichols.

ACKNOWLEDGMENT

The authors would like to thanks Control and Computer Systems Laboratory , School of Electrical Engineering and Informatics, Bandung Institute of Technology and Research Center For Electrical Power and Mechatronics-LIPI which has provided facilities in conducting this research.

REFERENCES

[1] Papic,Vladan.,Robot soccer,Olajnica 19/2,32000 Vukovar, Croatia,2010.
[2] Novak, Gregor., Springer,Richard., An Introduction to a Vision System used for a MiroSot Robot Soccer System, IEEE, pp. 101-108, 2004.
[3] Wong, Ching-Chang., Wang, Hoi-Yi., Li, Shih-An,. Fuzzy Controller Designed by GA for Two-wheeled Mobile Robots[J],International Journal of Fuzzy Systems, vol. 9, No. 1, pp. 22-30, March 2007.
[4] Gou-Jen Wang, C.T.Fong and Kang J.Chang, Neural network-based self-tuning PI controller for precise motion control of PMAC motors. IEEE Transactions on Industrial Electronics, 48, NO.2, pp. 408-416, April 2001.
[5] Wu,Yuzhi.,Wang,Haihua.,Application of Fuzzy self-tuning PID COntroller in Soccer Robot, IEEE, 8, 978-1-4244-3531-9,2008.
[6] Joko,Dwi PWA.,Warsito,Agung.,Triwiyatno,Aris,Penalaan Paramter Pengendali PID dengan algoritma genetik,Teknik Elektro, Undip.
[7] Coughanowr., R, Donald., Process Systems Analysis and Control, 2nd Edition, McGraw – Hill, New York, 1991.
[8] Ogata, Katsuhiko., Modern Control Engineering, Prentice Hall, New Delhi, 1984.
[9] Ogata, Katsuhiko., Teknik Kontrol Automatik (Sistem Pengaturan), Diterjemahkan oleh Edi Leksono, Jilid 1, Erlangga, Jakarta, 1995.
[10] Wang, Qing-Guo., Lee , Tong-Heng., Fung, Ho-Wang , Bi, Qiang., Yu Zhang., PID Tuning for Improved Performance, IEEE Transaction on Control System Technology, Vol. 7, No. 4, July 1999.
[11] ___________, YSR-A System Manual, Yujin Robotics Co., Ltd., 2003.
[12] Kim, J., Lecture Notes on Soccer Robotics. In ASEAN Technology Network for Multi-robot Cooperation System Development. Korea: Korea Advanced Institute of Science and Technology, 2003.
[13] Kim, Jong-Hwan., Kim, Dong-Han., Kim, Yong-Jae., Seow, Kiam-Tian., Soccer Robotics. In Springer Tracts in Advanced Robotics (Vol. 11). Berlin, Germany: Springer-Verlag, 2004.
[14] Firmansyah, Hanhan Husnah., Desain dan implementasi client-server pengolah data citra pada sistem robot soccer, Intitut teknologi Bandung, 2012.
[15] ASTROM, K. and HAGGLUND, T. , 1995, PID Controllers: Theory, Design, and Tuning, Instrument Society of America, ISBN 1-55617-516-7.
[16] OGATA, K..,Discrete-Time Control Systems, Prentice-Hall International, Inc.,ISBN 0-13-328642-8 , 1995.
[17] Guerra, R,S., Junior, Goncalves., Lages W.f., A digital PID controller Using RTAI, Federal University of Rio Grande do sul, Brazil.
[18] Barakbah, Ali Ridho., Reinforcement Learning paradigma baru dalam machine leraning, EEPIS-ITS, Surabaya.
[19] Qiang, Wang., Zhongli Zhan., Reinforcement learning model, Algorithms and its application, International conference mechatronic science, electro enginering and computer, August 19-22, China, 2011.
[20] Sutton,R,S., Barto,A., Reinforcement Learning: an Introduction , second printing, The MIT Press,1999.