

신경회로망을 이용한 쿼드로터의 자세 제어에 관한 연구

김성대*

A Study on the attitude control of the quadrotor using neural networks

Sung-Dea Kim*

요 약

최근 무인항공기(UAV : Unmanned Aerial Vehicle)에 대한 연구가 다양한 각도로 진행되어 군사용 비행체에 관한 연구에서 부터 민간용 비행체 및 일반 취미 활동용 비행체에 이르기까지 다양하게 연구가 진행되고 있다. 특히, 무인 소형 비행체에 대한 연구는 수직이착륙(VTOL : Vertical Take-Off and Landing)과 용이한 방향전환 및 정지비행(hovering)에 대하여 연구되고 있으며, 이러한 연구부분에 적합한 무인 소형 비행체가 쿼드로터(quadrotor)형 무인비행체를 중심으로 연구되고 있다. 이러한 무인 비행체에 대한 연구는 공기역학적 힘에 의해 부양되므로 복잡한 동역학 분석과정을 필요로 하고 있으며, 이러한 역학적 분석 및 실험적 모델을 바탕으로 제어를 설계하고 있다.

본 논문에서는 일반적인 PID 제어를 바탕으로 기본적인 자세제어를 구현한 후, 제어기 설계에 고려하지 못한 비선형적인 요소를 신경회로망(neural networks)의 강화학습(reinforcement learning) 알고리즘을 이용하여 일반적인 제어기 설계에 고려하지 못한 비선형적인 요소를 보완하여 보다 안정적인 쿼드로터의 자세제어 방안을 제시하고자 한다.

ABSTRACT

Recently, the studies of the Unmanned Aerial Vehicle(UAV) has been studied a variety from military aircraft to civilian aircraft and for general hobby activity aircraft. In particular, for small unmanned aircraft research for the ease of turning and hovering and Vertical-Off Take Landing(VTOL), have been studied mainly quadrotor unmanned aircraft is a type suitable for this study of small unmanned aircraft. The studies of these unmanned aircraft is the kinetic analysis requires complex processes, because these support by the aerodynamic forces on the unmanned aircraft study, and the controller design based on these dynamical analysis and experimental model analysis.

In this paper, after the implementation of the basic attitude control based on a general PID controller, we propose concept design of the attitude control method on quadrotor attitude control by using the reinforcement learning algorithm of neural networks for non-linear elements not considered in the controller design.

키워드

Quadrotor, Attitude Control, Neural Networks, Reinforcement Learning
쿼드로터, 자세제어, 신경회로망, 강화학습

* 교신저자(corresponding author) : 동명대학교 전기공학과(jksd@tu.ac.kr)

접수일자 : 2014. 07. 30

심사(수정)일자 : 2014. 08. 21

게재확정일자 : 2014. 09. 19

1. 서론

최근 무인비행체(UAV : Unmanned Aerial Vehicle)에 대한 연구가 군사용 및 산업용으로 활용하기 위해서 다양한 각도로 연구가 진행되고 있다. 특히, 산업분야에 활용하고자 하는 무인비행체에 관한 연구는 쿼드로터(quadrotor)형 비행체를 중심으로 연구가 진행되고 있다. 이렇게 쿼드로터에 대한 연구가 활발한 이유는 무인 소형 비행체가 가져야 할 기능적인 요소인 수직이착륙(VTOL : Vertical Take-Off and Landing)에 대한 기능과 무인 소형 비행체의 운전에 있어서 방향전환 및 방향제어 기능 및 필요에 의해 정지비행(hovering) 까지도 고려할 수 있어 연구 활용도가 높다고 할 수 있다.

쿼드로터에 대한 연구는 공기역학적 힘에 의해 부양되므로 복잡한 동역학적 해석과 복잡한 모델링 과정이 필요하다. 이러한 쿼드로터의 역학적 해석을 통한 선형화 작업을 거친 뒤 제어에 필요한 매개변수를 정의하여 반복적인 실험 또는 다양한 이론적 접근을 통하여 쿼드로터 시스템을 제어하고 있다. 이러한 기존의 선형적인 제어기 설계기법으로는 원하는 성능을 얻기가 힘들다. 이러한 단점을 보완하기 위하여 강인성과 적응성이 뛰어난 비선형 제어기인 슬라이딩 모드(SM : Sliding Mode) 제어, 신경회로망(NN : Neural Networks)을 이용한 제어, H_∞ 제어, 비주기적인 외란(disturbance)에 대하여 강인성을 확보할 수 있는 외란 관측기(DOB : Disturbance Observer)를 활용한 제어 등 다양한 비선형 제어기를 응용하여 쿼드로터 제어에 적용하고 있다[1-3].

인공지능(AI : Artificial Intelligent)의 학습방법 중 강화학습(reinforcement learning)은 동적 프로그래밍(DP : Dynamic Programming) 방식과 교사학습(supervised learning) 방식이 혼합된 방식으로 학습을 진행하는 에이전트(agent)와 에이전트 외부에 존재하는 환경(environment)에 대하여 반복적으로 시행착오를 거쳐 내부 상호작용(interacting)을 학습한다. 이러한 시행착오를 통하여 각 상태에서의 최적 행동(action)을 결정하는 학습 알고리즘이다. 이러한 특징 때문에 기구동역학 시스템의 모델링 과정을 거치지 않고 주어진 환경에서 행동을 시도하고, 시도한 행동에 대한 결과를 평가하여 강화값을 결정하므로 시스

템 모델에 덜 의존하고 있는 중요한 특성도 가지고 있다. 이러한 강화학습 중 가장 일반적으로 방법이 Q-learning이며, 학습방법 또한 주어진 환경과 상호작용을 통하여 보상(reward)을 최대화 하는 제어정책을 학습하는 방법이다. 때문에 비선형적인 요소가 많은 복잡한 시스템에 적용이 가능하고, 시스템 모델을 알지 못하는 상황에서 온라인(on-line) 학습이 가능하다[4-5]. 즉, 시스템에 대한 정확한 사전지식이 없어도 학습이 가능하고, 적응성을 보장하기 때문에 로봇 또는 기계학습에 유용하게 사용할 수 있는 알고리즘이다[6-7].

본 논문에서는 일반적인 PID 제어기를 바탕으로 쿼드로터에 대한 기본적인 자세제어를 Matlab Simulink를 통하여 시뮬레이션 하였다. 또한, 신경회로망 기반의 강화학습을 통하여 일반적인 제어기 설계에 고려하지 못한 비선형적인 요소를 보완하여 제어성능을 향상시키는 방안을 제시하고자 한다.

II. 본론

2.1. 쿼드로터 시스템 구성

본 논문에서 사용된 쿼드로터의 구성은 가장 일반적인 십자형 프레임 구조로 하고 각각의 끝부분에 4개의 로터로 구성된다. 각각 마주보는 2개의 로터는 같은 방향으로 회전하도록 하고, 이웃하는 로터는 반대방향으로 회전하게 하여 로터의 회전운동에서 발생하는 반발 토크를 서로 상쇄하도록 하여 균형을 잡도록 한다. 이러한 쿼드로터 시스템을 그림 1과 같다.

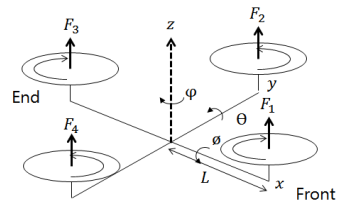


그림 1. 쿼드로터 모델
Fig. 1 Quadrotor model

4개의 로터 회전력에 의하여 추력(F)을 발생하여 쿼드로터를 원하는 위치로 움직일 수 있다. $F_1 \sim F_4$ 의

추력에 의해 쿼드로터의 상승과 하강이 이루어지고, F_2 와 F_4 의 차이에 의해 롤(roll) 관성모멘트(moment)가 발생하고, F_1 과 F_3 의 차이에 의해 피치(θ ,pitch) 관성모멘트가 발생한다. 요(ψ ,yaw)에 대한 관성모멘트는 $(F_1 + F_3 - F_2 - F_4)$ 의 합력에 의해 운동방향이 결정된다. 이러한 추력의 상관관계에 의해 쿼드로터의 운동방정식이 결정되고 쿼드로터의 제어에 활용된다.

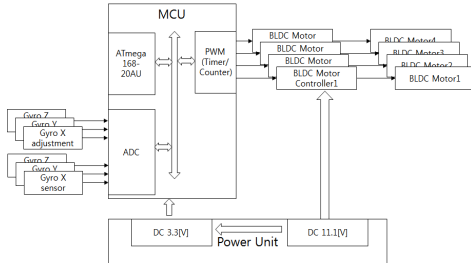


그림 2. 쿼드로터 제어보드 구성
Fig. 2 Configuration of quadrotor control board

2.2. 쿼드로터 시스템 모델링

그림 1을 바탕으로 쿼드로터 시스템의 운동방정식은 직각 좌표계 x, y, z의 3축에 대하여 각각 롤, 피치, 요로 정의하고 식(1)로 표현할 수 있다.

$$\begin{aligned} m\ddot{x} &= F(\sin\phi \sin\psi + \cos\phi \sin\theta \cos\psi) \\ m\ddot{y} &= F(\cos\phi \sin\theta \sin\psi - \sin\phi \cos\psi) \\ m\ddot{z} &= F(\cos\phi \cos\theta) - mg \end{aligned} \quad (1)$$

$$\begin{aligned} F &= F_1 + F_2 + F_3 + F_4 \\ \ddot{\phi} &= L(F_2 - F_4)/J_{xx} \\ \ddot{\theta} &= L(F_1 - F_3)/J_{yy} \\ \ddot{\psi} &= C(F_1 - F_2 + F_3 - F_4)/J_{zz} \end{aligned} \quad (2)$$

식(1)은 원심력과 구심력을 제외한 단순화한 쿼드로터의 운동방정식이다. 식(2)의 F 는 각 로터의 추력의 합이며, J_{xx}, J_{yy}, J_{zz} 는 x, y, z의 3축에 대한 관성모멘트이다. L 은 로터와 쿼드로터 중심과의 거리이며, C 는 로터의 회전에 의해 발생하는 반발토크에 대한 변환상수이며, g 는 중력가속도이다.[3] 로터에 존재하는 추력의 변화에 의하여 쿼드로터에 발행하는 관성모멘트와의 상관관계를 식(2)에 제시하였다[2].

2.3. 강화학습

강화학습의 특징은 목표 지향적으로 학습이 진행되며, 현재의 시행착오와 주어진 환경과의 상호작용을 통하여 학습이 이루어지는 알고리즘이다. 즉, 시스템이 처해진 환경에서 에이전트가 보상을 최대화 할 수 있는 행동전략을 학습하는 것이라 할 수 있다. 이러한 에이전트 중심의 행위와 보상에 의해 현재 환경에 대한 보상을 최대로 할 수 있는 행동을 선택하는 것이다. 강화학습이 이루어지는 시간 t 가 무한하다고 가정하고, 에이전트에게 입력된 보상값을 r_{t+n} 라 두고, 보상값의 합을 R_t 라 하면 보상함수는 식(3)과 같이 표현된다.

$$R_t = r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \dots = \sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \quad (3)$$

γ 는 할인율(discount rate)은 $0 < \gamma < 1$ 값을 가지도록 하여 보상값의 합이 발산하지 않도록 해야 한다 [4].

강화학습은 상태(state)와 행동의 쌍으로 이루어진다. 상태에 해당하는 상태가치함수(state value function) $V^\pi(s)$ 는 상태 s_t 에서 정책(policy) π 에 의한 보상값 r_t 의 합으로 주어지며, 행동에 해당되는 상태 행동가치함수(state action value function) $Q^\pi(s, a)$ 는 상태 s_t 에서 행동 a_t 와 정책 π 에 의한 강화값 r_t 의 합으로 나타낸다. 이러한 정의를 식 (4)~(6)에 표현하였으며, 식 (6)의 a' 은 다음상태의 행동을 의미한다.

$$\begin{aligned} V^\pi(s) &= E_\pi \{ R_t | s_t = s \} \\ &= E_\pi \left\{ \sum_{k=0}^{\infty} \gamma^k r_{t+k+1} | s_t = s \right\} \end{aligned} \quad (4)$$

$$\begin{aligned} Q^\pi(s, a) &= E_\pi \{ R_t | s_t = s, a_t = a \} \\ &= E_\pi \left\{ \sum_{k=0}^{\infty} \gamma^k r_{t+k+1} | s_t = s, a_t = a \right\} \end{aligned} \quad (5)$$

$$\begin{aligned} V^*(s) &= \max_{\pi} V^\pi(s) \\ Q^*(s, a) &= \max_{\pi} Q^\pi(s, a) \\ \pi^*(s) &= \arg \max_{a \in A} Q^*(s, a') \end{aligned} \quad (6)$$

강화학습의 진행과정은 목적을 만족하는 상태와 행동의 쌍에는 보상값을 크게 할당하고, 목적을 달성하지 못한 쌍에는 보상값은 낮추게 하여 최종적으로 최적의 행동 전략을 찾아내는 것이라 할 수 있다. 이러한 보상을 통해 최적화된 상태와 행동을 결정하는 알고리즘 중 하나가 Q-learning 이며, Q-learning의 가치함수 갱신방법은 식(7)과 같다.

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [R_{t+1} + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)] \quad (7)$$

식(7)의 α 는 학습률(learning rate)을 의미한다.

본 논문에서는 쿼드로터의 자세 제어를 위하여 자세 오차를 최소화하기 위하여 쿼드로터의 상태와 제어 행동에 대한 평가를 통하여 최적의 제어입력을 도출하는데 있다고 할 수 있다.

III. 시스템 구성과 시뮬레이션 및 실험 결과

3.1 시스템 구성

본 논문에 사용된 쿼드로터 메인보드의 구성은 추력을 제어하기 위한 PWM 출력 4개와 쿼드로터의 자세 측정을 위한 자이로(gyro) 센서 2개(로, 피치-1개, 요-1개), 자세에 대한 오프셋 보정을 위한 보정용 가변저항 3개에 대한 5개의 아날로그 입력으로 구성된다. 추가로 쿼드로터의 내부 상태를 모니터링하기 위하여 USART 포트 1개로 구성된다. 이러한 시스템의 구성을 그림 3에 제시하였다.

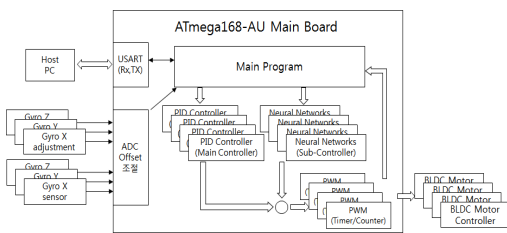


그림 3. 쿼드로터 시스템 블록선도
Fig. 3 Block diagram of quadrotor system

본 논문에 사용된 쿼드로터는 <http://www.multicopter.kr>의 플라이캠(FlyCam) 멀피콥터 모델을 사용하였다. 사용된 MCU는 ATmega168-20AU이며, 개발환

경은 Atmel Studio 6.2 기반으로 하였다. 쿼드로터의 자세를 계측하기 위한 자이로 센서 2개(Gyro IDZ-500(Yaw), Gyro IDG-500(Roll, Pitch))와 충전 가능한 3cell DC 11.1[V] Li-PO 배터리와 제어보드 및 자이로센스의 전원 공급을 위한 DC/DC 컨버터 IL1117S-3.3을 이용하여 DC 3.3[V]를 공급하도록 되어있으며, 쿼드로터의 추력을 3.2[Kg] 이상의 추력을 얻고, 고효율의 FlyCam925 BLDC(Brushless DC Motor) 모터 4조와 10x4.5[inch] 프로펠러 4개, 각각의 BLDC 모터의 속도 제어용 전자속도 제어기(ESC, Electronic Speed Controller) 4개를 조합하여 쿼드로터 시스템을 구성하였다. 쿼드로터 시스템의 구성은 아래의 그림 4로 제시하였으며, 실험의 안정성을 고려하여 z축에 해당하는 높이 파라메타는 고정하였으며, 롤(F_1, F_3) 또는 피치(F_2, F_4)에 해당하는 로터 쌍을 고정하고, 나머지 2개의 로터에 대한 제어를 통하여 자세제어가 이루어지도록 실험 환경을 구성하였다.



그림 4. 쿼드로터 실험환경
Fig. 4 Quadrotor test environment

3.2 PID 기반의 시뮬레이션

쿼드로터 상태방정식은 뉴턴방정식을 기초로 하였으며, 그림 1의 파라메타를 기준으로 식(1), 식(2)의 결합으로 이루어진다. 여기에서 모터에 대한 상수 및 기타 외란은 생략하였다.

상승/하강 운동은 수직축 z축 따라 움직이며 식(8)과 같다.

$$m\ddot{z} = (F_1 + F_2 + F_3 + F_4) - mg \quad (8)$$

진전/후진에 대한 역학적 운동은 식(9)와 같이 x축에 연결된 피치(θ)각의 크기에 따라 결정된다.

$$\begin{aligned} m\ddot{x} &= (F_1 + F_2 + F_3 + F_4)\sin\theta \\ \ddot{\theta} &= L(F_1 - F_3)/J_{yy} \end{aligned} \quad (9)$$

왼쪽/오른쪽에 대한 역학적 운동은 식(9)와 같이 y 축에 연결된 물(ϕ)각의 크기에 따라 결정된다.

$$\begin{aligned} m\ddot{y} &= (F_1 + F_2 + F_3 + F_4)\cos\phi \\ \ddot{\phi} &= L(F_2 - F_4)/J_{xx} \end{aligned} \quad (10)$$

요(ψ)의 크기는 식(11)과 같이 대각선 합으로 나난다.

$$\ddot{\psi} = C(F_1 - F_2 + F_3 - F_4)/J_{zz} \quad (11)$$

식(8)~식(11)을 결합하여 그림 5와 같이 Matlab Simulink 블록 다이어그램으로 표현하였으며, 임의의 PID 제어기를 설계하여 시뮬레이션을 수행하였다.

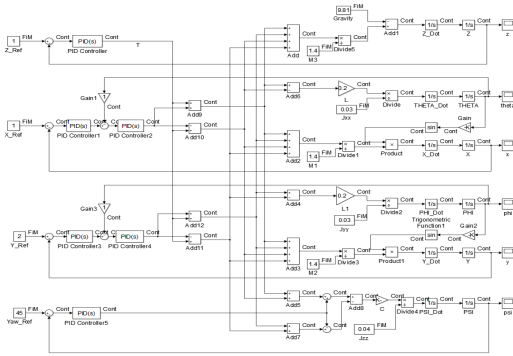


그림 5. PID 기반의 쿼드로터 모델에 대한 블록 다이어그램

Fig. 5 Block diagram of quadrotor model based on PID controller

시뮬레이션을 위한 적절한 PID 계수는 시행착오를 거쳐 표 1과 같이 선정하였다.

표 1. PID 제어기 파라메타
Table 1. PID controller parameter

	x-axis	y-axis	z-axis	Yaw(ψ)
P	4	10	0.1	5
I	0.05	0.9	25	2
D	5	10	20	2

그림 6~그림 8은 시뮬레이션 조건에 따른 결과를 제시하였다.

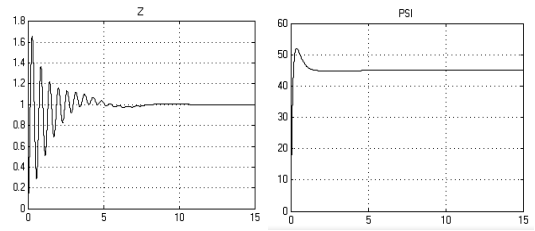


그림 6. z_Ref=1, $\psi=45$ 에 대한 시뮬레이션 결과(x, y=0)
Fig. 6 Simulation results for z_Ref=1, $\psi=45(x=0, y=0)$

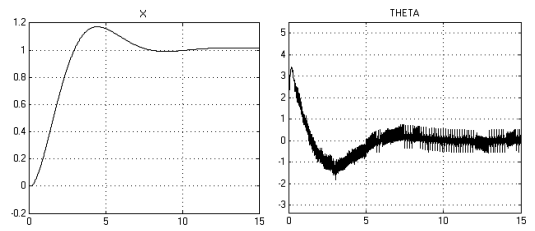


그림 7. x_Ref=1에 대한 시뮬레이션 결과(y,z=0, $\psi=0$)
Fig. 7 Simulation results for x_Ref=1(y,z=0, $\psi=0$)

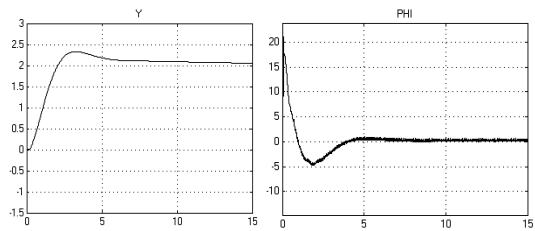


그림 8. y_Ref=2에 대한 시뮬레이션 결과(x, z=0, $\psi=0$)
Fig. 8 Simulation results for y_Ref=2(x, z=0, $\psi=0$)

3.3 강화학습을 통한 쿼드로터 실험

쿼드로터의 운동성능과 제어성능을 확인하기 위하여 dSPACE 사의 Micro-AutoBox(ds1401)을 이용하여 실험 시스템을 구성하였다. 실험 조건은 물과 피치에 해당하는 F_1, F_3 로터는 고정하고, F_2, F_4 에 해당하는 로터를 제어하여 운동 및 제어성능을 분석하였다.

로터제어에 대한 PWM 펄스 주파수는 내부클럭 1[Mhz]에서 timer1(16bit)을 내부 타이머를 이용하여 50[Hz]로 설정하였다.

실험은 그림 3과 같이 주제어기는 offset을 조절할 수 있는 자리로 센스 기반의 PID 제어기와 임의의 외란과 PID 제어기의 성능을 보상할 수 있는 강화학습

기반의 보조제어를 통하여 로터 2개에 대한 PWM 듀티비를 보상하여 운동성능에 대하여 실험하였다.

그림 9는 임의의 PID 제어를 통해 제어되는 상태이며, 수평상태가 유지된 상태에서 외란이 발생하였을 때 목표치인 엔코더 값에 대하여 쿼드로터의 움직임 수렴해 가지 못하는 것을 확인할 수 있다.

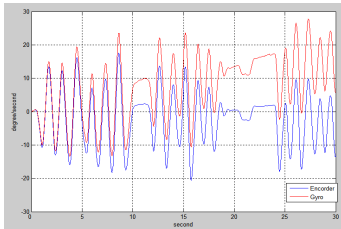


그림 9. 주제어기(PID)에 의한 쿼드로터의 운동특성($F_1 = F_3 = 0$)

Fig. 9 Movement characteristics of the quadrotor by the main controller(PID).($F_1 = F_3 = 0$)

그림 10은 주제어기와 강화학습을 통한 보조제어를 통하여 제어되었을 때 목표치인 엔코더 값에 대하여 쿼드로터의 움직임이 추종해 나가고 있는 것을 확인할 수 있다.

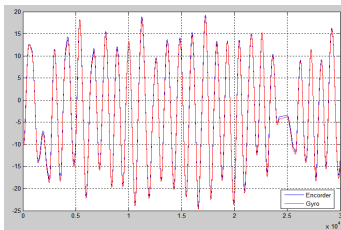


그림 10. 주제어기(PID)+보조제어기(강화학습)을 통한 쿼드로터 운동특성($F_1 = F_3 = 0$)

Fig. 10 The characteristics of quadrotor movement through the main controller(PID) + secondary controller(reinforcement learning).($F_1 = F_3 = 0$)

그림 11은 하나의 축에 대한 실험을 거친 후 주제어기를 우선으로 하고, 보조제어기의 보조역할을 추가하여 운전된 상태를 그림 11에 제시하였다,



그림 11. 쿼드로터의 운전
Fig. 11 Operation of the quadrotor

V. 결 론

쿼드로터의 운동성능은 강화학습의 초기 상태와 행동을 결정에 대한 정보저장 테이블의 불완전한 상태로 인하여 초기 제어성능은 오히려 기존의 PID 제어기 보다 불안정하였으며, 일정한 시간이 경과한 후에 서야 안정된 제어성능을 보인 것을 확인할 수 있었다.

향후 연구과제는 강화학습을 통하여 더욱 빠르게 안정화 시킬 수 있는 방법을 연구하고, 실제 쿼드로터의 운전상태와 강화학습의 정보저장 테이블을 비교 분석하여 더욱 빠르고 안정적인 학습방법을 연구하고자 한다.

감사의 글

이 논문은 2012학년도 동명대학교 교내학술연구비 지원에 의하여 연구되었음

참고 문헌

- [1] I. Maza and A. Ollero, "Multiple UAV Cooperative Searching Operation using Polygon Area Decomposition and Efficient Coverage Algorithms," In *Proc. the 7th Int. Symp. on Distributed Autonomous Robotic Systems*, Toulouse, France, 2004, pp. 211-220.
- [2] S. Jeong, S. Jung, and M. Tomizuka, "Attitude control of a quad-rotor system using an acceleration-based disturbance observer : An empirical approach," *The 2012 IEEE/ASME Int. Conf. on Advanced Intelligent Mechatronics*, Kao-

- hsiung, Taiwan, July 2012, pp. 916-921.
- [3] J. Kim, M. Kang, W. Chung, H. Kim, and I. Suh, "On the robustness control of a quadrotor VTOL aircraft," *J. of Institute of Control, Robotics and Systems*, vol. 14, no. 12, 2008, pp. 1260-1265.
- [4] J. Choi, "Voiced-Unvoiced-Silence Detection Algorithm using Perceptron Neural Network," *J. of the Korea Institute of Electronic Communication Sciences*, vol. 6, no. 2, 2011, pp. 237-242.
- [5] J. Choi, "Speech and Noise Recognition System by Neural Network," *J. of the Korea Institute of Electronic Communication Sciences*, vol. 5, no. 4, 2010, pp. 237-242.
- [6] L.-J. Lin, *Reinforcement Learning for Robots Using Neural Networks*. Pittsburgh, PA, Carnegie-Mellon Univ., 1993.
- [7] M. Hoffmann, J. Jang, and J. Tomlin, "Multi-Agent Quadrotor Tested Control Design : Integral Sliding Mode vs. Reinforcement Learning," *IEEE/RSJ Int. Conf. on Intelligent Robot and Systems*, Alberta, Canada, Aug. 2005, pp. 3712-3717.

저자 소개



김성대(Sung-Dea Kim)

1984년 동아대학교 물리학과 졸업
(이학사)

1986년 동아대학교 대학원 전자공
학과 졸업(공학석사)

1996년 동아대학교 대학원 전자공학과 졸업(공학박사)

1991년~현재 동명대학교 전기공학과 부교수

※ 관심분야 : 자동제어, 인공지능