

RF 통신을 통한 데이터 모니터링 및 강화학습을 활용한 비행 시뮬레이터에 관한 연구

김윤수^a · 이민형^b · 신나라^c · 이승호^c · 이기택^{c*}

LAUNCH DATA MONITORING SYSTEM AND LAUNCH SIMULATOR BY USING REINFORCEMENT LEARNING FOR FLIGHT CONTROL

Yoonsoo Kim^a · Minhyung Lee^b · Nara Shin^c · Seungho Lee^c · Gitaek Lee^{c*}

^a*School of Computer Engineering, Gachon University, Korea*

^b*School of AI-Software, Gachon University, Korea*

^c*School of Mechanical Engineering, Gachon University, Korea*

**Corresponding author. E-mail: ximons22@gachon.ac.kr*

ABSTRACT

This study aims to capture real-time video, axis, and altitude data by wireless communication. This could monitor the status of the rocket system and capture useful data. For stable communication, we will do Proportional-Integral-Differential(PID) control with the Reaction wheel to control roll 축 rotation. In addition, to increase the stability of rocket flight, we made our own simulator to test rocket stabilization by Thrust-Vector-Control(TVC) without fins. In this case, we use AI model with Reinforcement learning. We consider this simulator can be used in the next rocket to control the posture of the rocket.

초 록

본 논문은 로켓 발사 시 수집할 수 있는 영상, 자세, 고도 등의 데이터를 통신으로 얻고 이를 로켓의 실시간 상태 확인 및 주요 데이터 수집으로 활용한 연구이다. 안정적인 데이터 수집을 위해 Proportional-Integral-Differential(PID) control을 사용했고 로켓의 roll - axis 회전 제어를 Reaction wheel을 통해 수행한다.

추가로 로켓의 안정적인 발사를 위해 Reinforcement learning을 통한 Thrust-Vector-Control(TVC)를 자체 시뮬레이션 내에서 구현하여 향후 발사에 적용할 수 있도록 하였다. 이는 22년 발사대회에 사용할 실제 엔진 추력 데이터를 기반으로 구성되었다.

Key Words: PID control(PID 제어), Reaction wheel(리액션 휠), Reinforcement learning (강화학습), TVC(추력 편향 제어)

Nomenclature

D_t : Nozzle Throat Diameter
 D_e : Nozzle Exit Diameter
 F : Thrust
 γ : Specific Heat Ratio
 P_a : Chamber Average pressure
 P_M : Chamber Max Pressure
 P_e : Nozzle Exit pressure
 ϵ : Nozzle Expansion Ratio
 KNSB : Potassium Nitrate-Sorbitol

1. 서 론

본 논문의 목적은 KNSB 연료를 사용한 모델 로켓을 제작하여 다양한 정보를 수집, 가공하여 활용하는 것에 있다. 로켓에 탑재한 여러 장비를 이용하여 영상, 고도, 자세, 온·습도 정보를 수집한다.

로켓의 정보 수집 현황과 상태를 확인하기 위해서 영상 정보를 제외한 고도, 자세, 온·습도 정보는 실시간 Radio frequency 통신을 통해 지상국으로 데이터를 송신한다. 영상 정보의 경우, 송신 속도 문제가 있으므로 기본적으로는 SD 카드에 영상을 저장하고 WIFI 통신이 가능한 200M 거리 내에서 데이터 수신할 예정이다. 안정적인 데이터 송·수신을 위해 로켓의 roll 축 회전을 최소화하는 것이 필요하다. 이를 위해 Proportional-Integral-Differential(PID) control을 이용한 Reaction wheel을 사용한다.

수집한 정보를 바탕으로 실제 이론값과의 비교 분석을 통해 로켓의 효율과 안정성을 높이는 것이 최종 목표이다. 최소 3kg 이상의 추진체를 상공 300m 이상까지 상승시키기 위해 I급 이상 엔진을 개발하였다.

또한 Thrust-Measurement-System(TMS) 결과를 바탕으로 Thrust-Vector-Control(TVC)을 위한 자체 시

뮬레이터 개발을 진행하였다. 해당 시뮬레이터는 Reinforce learning을 통해 실제 상황을 가정한 가상 공간에서 PID control과 유사하게 구현된다. 이는 향후 TVC 장치 개발에 유용한 사전 자료로 활용될 것으로 기대된다.

2. 본 론

2.1 추진

설정된 목표를 달성하기 위해 안정적인 성능의 추진기관이 필요하다. 이를 위해 다음과 같이 추진기관의 설계와 실험을 진행하였다. 설계에 앞서 그레인과 챔버의 제원을 Table 1과 같이 설정하였다. 설정된 값들은 METEOR 시뮬레이터에서 탑재 가능한 엔진 크기이며 I급 이상 성능을 고려한 것이다.

Table 1. KNSB ENGINE VALUE

CCD* (mm)	NTD* (mm)	Chamber length (mm)	Grain core diameter (mm)	Grain mass (g)
40	10	210	15	373

*Combustion Chamber Diameter **Throat Diameter

2.1.1 연소실 설계

보수적인 설계를 통해 일반적인 고체 추진체의 안전계수[1]의 약 2배 이상으로 설정하였다. METEOR 시뮬레이터에서 도출된 압력 값은 Fig. 1이고 사용한 AL 6061 챔버의 제원은 Table 2와 같다. 이를 근거로 Eq. 1을 이용하여 도출된 안전계수는 최대 내압 기준으로 6.34이다.

Table 2. ENGINE CHAMBER SET VALUE

Inner diameter (mm)	Outer diameter (mm)	Thickness of chamber (mm)	yield strength of AL 6061 [2]
40	50	5	110MPa

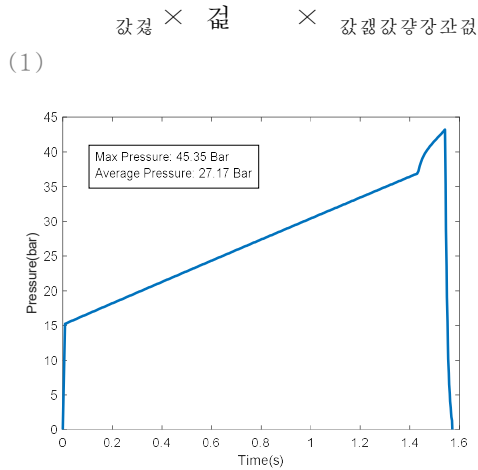


Fig. 1 pressure result in meteor simulator.

2.1.2 노즐 설계

노즐 설계는 Fig. 1에서 도출된 평균 내압에 근거하였다. 노즐 최적 확장비를 구하는 Eq. 2와 Table 1의 값을 고려하여 노즐 팽창비를 설정하고 노즐을 설계하였다. Eq. 2를 통해 구한 노즐의 노즐 목과 노즐 출구 면적의 최적 비율은 5.86이다. 일반적인 Conical nozzle 설계에 따라 입구 각을 30° , 출구 각을 13° 로 설정하였고, 이는 Fig. 2와 같다[2]. 노즐의 소재는 연소 중 노즐 목의 지름 확장을 방지하기 위해 STS 304를 사용하였다.

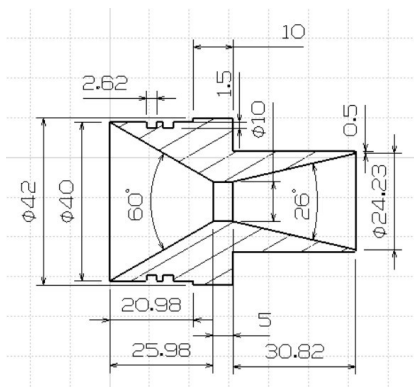


Fig. 2 rocket nozzle design.

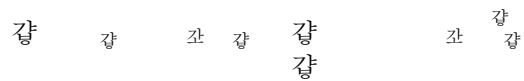
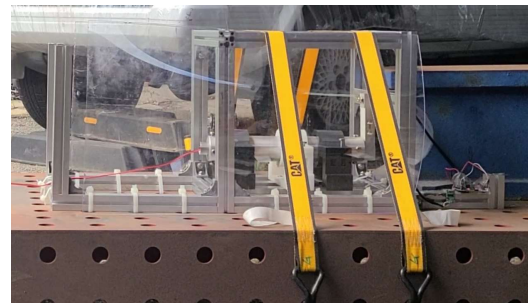


Fig. 3 TMS design.

2.1.3 TMS

TMS 장치 구성은 Fig. 3과 같다. 장치의 프레임은 알루미늄 중량 프로파일로 구성하였다. 엔진 고정장치를 통해 정해진 축 상에서만 엔진이 움직이도록 하였고, 장치의 사출 방향을 제외한 모든 면은 3T 폴리카보네이트 판을 결합하여 안전사고에 대비하였다. 측정 장비는 HX711 앰프를 결합한 막대형 로드셀을 사용하였고, 80Hz로 측정하였다.



2.1.4 TMS 결과

TMS 진행 결과와 METEOR 시뮬레이터에 의한 이론값의 오차는 Table 3과 같다. 연소 후 STS 304 노즐의 질량 차는 사용한 저울의 최소 측정 범위인 1g 단위 이하이고, D_e 또한 Fig. 4와 같이 원형을 유지하고 있다. 일반적으로 사용하는 알루미늄 6061 노즐과 비교했을 때 내삭마성이 매우 강한 것을 확인하였다. 실제로 Fig. 5와 같이 추진 영상 분석에서 불꽃 등의 금속 연소 현상이 발생하지 않았다. 또한 노즐과 벌크 헤드의 챔버 결합 부분에서 연기 등의 압력 유출 징후는 발견되지 않았다. 측정된 추력 그래프는 Fig. 6과 같다.



Fig. 4 after combustion nozzle.



Fig. 5 TMS at max thrust.

2.1.5 TMS 결과 분석

TMS로부터 얻은 추력 값을 바탕으로 최대 추력, 총충격량, 비추력을 Eq. 3을 바탕으로 구하였다. 또한 METEOR 시뮬레이터와의 비교를 Table 3과 같이 나타내었다. Table 3과 Fig. 6을 통해 METEOR 시뮬레이터와 오차를 확인할 수 있다. 오차를 발생시키는 노즐의 삭마, 외부로의 압력 유출 등이 관찰되지 않았으나 상당한 오차가 발생하였다. 따라서 노즐의 삭마와 압력 유출을 제외한 결과에 영향을 줄 외부 요인을 ① 로드셀 캘리브레이션, ② 연소 부산물의 노즐 흡착, ③ 그레인의 수분 흡수로 추정하였다.

Table 3. METEOR SIMULATOR VS TMS RESULT

	Meteor Simulator	TMS result	Relative error(%)
M.T* (N)	480.95	403.87	-16.02
I _t ** (N·s)	450.29	494.82	9.89
I _s *** (sec)	122.83	135.22	10.09
B.T**** (sec)	1.57	2.1	33.76%

*Max Thrust **Total Impulse ***Specific Impulse ****Burn Time

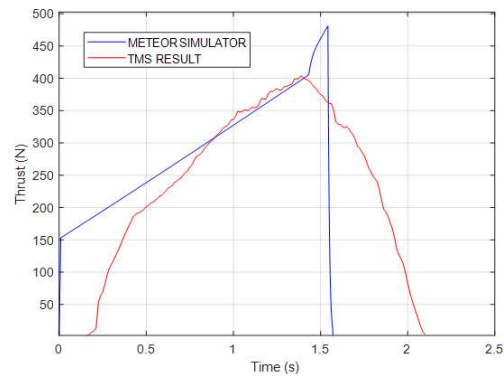


Fig. 6 meteor simulator vs TMS result.

$$\text{결과} = \text{결과} \times \frac{\text{중량}}{\text{중량}} \times \frac{\text{중량}}{\text{중량}}$$

(3)

연소 이후 챔버를 조사하여 모든 연료가 연소된 것을 확인하였다. 시뮬레이터와 같은 중량의 연료가 연소 되었지만 10%의 오차가 발생한 것은 로드셀이 정확한 값을 얻지 못한 것, 즉 캘리브레이션에서 발생한 오차라고 해석할 수 있다. 실제 로드셀 캘리브레이션을 최대 추력에서 받게 되는 중량을 기준으로 실험하기엔 제약이 많다. 이 점을 보완하기 위해 추후 실험에서는 특정 중량에서 얻은 캘리브레이션 값이 전체 측정 범위에서도 높은 신뢰성을 갖는 제품을 사용하여 해당 문제점을 해결할 수 있을 것으로 기대한다.

Fig. 6 약 1.0초 ~ 1.5초 구간에서의 값이 다른 구간에 비해 비선형적인 것을 확인할 수 있다. 연소 이후, Fig. 7과 같이 연소 부산물이 노즐의 입·출구에 흡착된 것을 확인하였고, 그것을 해당 원인으로 추측하였다. KNSB의 연소 부산물은 수산화칼륨(KOH)[3]으로 보이나 정확한 성분 분석이 필요하다. 또한 연소 상황을 고려해 부산물이 노즐에 흡착된 이유를 찾고, 이를 제거하는 방안을 모색하는 것이 추후 연소 효율을 증가시킬 것으로 보인다.

마지막으로 TMS 당시 연소 시험장의 습도가 90% 이상이었던 점, 점화를 2회 시도하는 과정에서 약 5분간 연료가 습한 상황에 노출된 점이

연소시간 증가에 큰 영향을 끼친 것으로 추정한다. 습도에 취약한 KNSB 연료인 만큼, 추후 TMS에서는 점화 실패 시 예비 연료로 대체하는 등의 방안으로 습기에 노출되는 상황을 최소화하여 해당 문제를 해결할 수 있을 것으로 기대한다.

이에 더하여 시뮬레이터와 더 유사한 추진을 위해선 목표 내부 압력 도달 전까지 챔버의 압력을 유지 시키는 마개가 고안되어야 할 것으로 보인다. 이 부분은 추후 추가 TMS에서 보완 실험을 통해 적정한 형상의 노즐 마개를 고안할 예정이다.



Fig. 7 nozzle after TMS.

2.2 동체

동체는 낙하산과 낙하산 사출장치, 내부 장비 고정 프레임, 핀, 노즈콘으로 구성되어 있으며, 안정적인 비행을 돕고 외부 충격으로부터 내부 장비들을 보호할 수 있도록 설계되었다.

2.2.1 노즈콘

노즈콘은 추진체가 안정적으로 공기를 가르며 상승하게 하는 중요한 요소이다. 노즈콘 형상은 발사체의 속도에 근거하여 선정된다. 따라서 대략적인 발사체의 비행 속도를 파악하기 위해 현재 목표하고 있는 로켓의 상승 시 무게인 2.5kg과 TMS 결괏값을 사용하여 OPEN ROCKET 프로그램에서 Fig. 8과 같이 값을 얻었다. 최대 높이는 400m, 최고 속도는 0.559 Mach로 190m/s이다. 즉, 최대 아음속으로 상승한다는 사실을 알 수 있다.

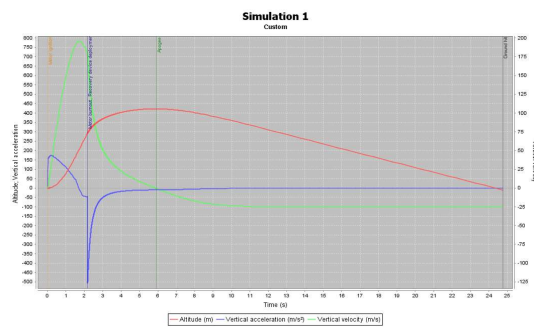


Fig. 8 open rocket simulator test result.

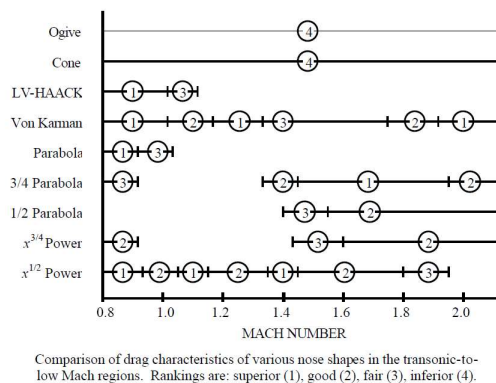


Fig. 9 general shapes of nose cone.

해당 결과와 Fig. 9의 값을 토대로 노즈콘을 설계하였다. Fig. 9는 ①에 가까울수록 항력계수가 작다는 것을 의미한다[4]. 이를 바탕으로 Von Karman 형상을 선택하였다. Von Karman 형상은 항력계수가 0.23으로 매우 낮다. 이는 Haack Series 노즈콘에서 기울기를 정하는 계수인 C가 0인 경우이기도 하다. Haack Series 노즈콘은 Eq. 3과 Eq. 4를 이용하여 구한다. 여기서 Haack Series는 C가 0이므로 최종 Von Karman 노즈콘은 Eq. 5를 이용해서 설계할 수 있고, 이는 Fig. 10과 같다. 재질은 무게 및 내구성을 고려하여 PLA 3D 프린터 출력을 할 예정이다.

$$\text{arc}\text{검검} \quad \text{경} \quad (3)$$

낙하산이 강한 바람에 뒤집히는 것을 방지하기 위해 spill hole을 추가하였다. 이는 낙하산 면적의 약 4%로 설정하였다. 대략적인 낙하산 면적을 구하기 위해 Table 4와 같이 각 값을 설정하였고 이에 대한 낙하산 면적 추정치는 Table 5와 같다. 추정치는 노즈콘과 같이 OPEN ROCKET 프로그램으로 예상 발사체 세부 수치와 TMS 결과를 이용하여 구하였다. 실제 낙하산의 세부 수치는 실물 발사체의 세부 수치를 파악한 이후에 계산할 수 있을 것으로 기대한다.

Table 4. SET VALUE FOR PARACHUTE DESIGN

Descent Rate (강)	3m/s
Parachute Drag Coefficient (강)	1.17
Rocket Recovery Mass (강)	2.5kg
spill hole Diameter Percentage (강결강)	20%

Table 5. ESTIMATE OF PARACHUTE DESIGN

Parachute Reference Area (결)	3.95 m ²
Parachute Diameter (강결)	2.24 m
Spill Hole Area (결결강강)	0.152 m ²
Spill Hole Diameter (강결결강강)	0.44 m

2.2.3 내부 장비 고정장치

내부 장비 고정장치는 내부에 탑재되는 여러 전자 장비와 엔진을 고정 및 보호할 수 있도록 설계하였다. 탑재 장비가 달라지더라도 간단하게 부품 교체가 가능한 형태로 설계하였으며 조립 시간을 단축하기 위해 동체의 부분별 연결 부위의 결합을 Fig. 13과 같이 돌려서 결합할 수 있게 설계하였다.

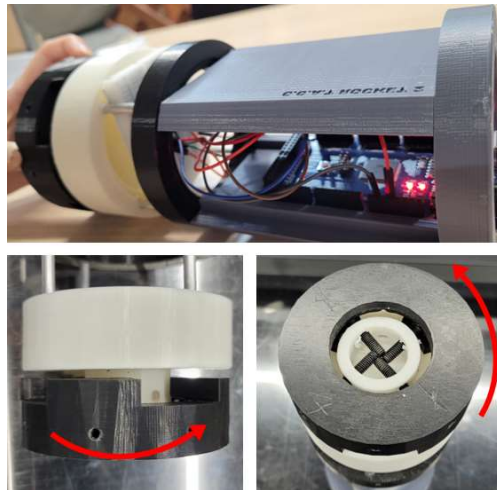


Fig. 13 inner mount and coupler.

2.2.4 핀

핀은 발사체의 압력 중심을 설정해 stability 값을 결정할 수 있게 한다. 따라서 탑재 장비와 위치에 따라 설계를 다르게 해야 한다. 이에 대하여 실측 무게 중심에 따라 핀을 결합할 수 있도록 Fig. 14과 같이 핀 마운트와 핀을 설계하였다. 핀 표면 가공과 내구성을 고려하여 ABS 재질로 3D 출력 후, 아세톤을 이용하여 후가공하였다. 실제 핀 설계는 모든 장비를 탑재한 동체의 무게 중심을 실측한 뒤, OPEN ROCKET 프로그램을 통해 핀의 형상을 결정한 뒤 장착한다. stability 값이 2가 되는 핀 형상을 사용할 예정이다.

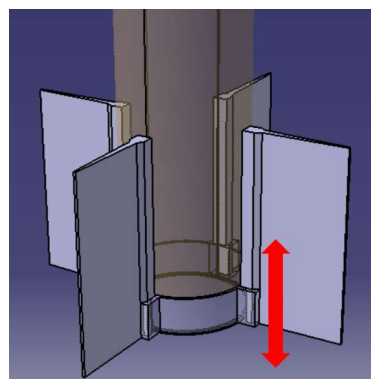


Fig. 14 fin mount and fins

2.3 데이터 수신

이번 발사체는 영상, 고도, 자세, 기압, 온도 측정을 목표로 한다. 해당 데이터 처리를 위해 고성능의 라즈베리 파이 4를 지상국과 발사체에 탑재한다. IMU 센서인 BNO055를 이용하여 자세 데이터를 수신받고, BMP 280을 이용해서 고도, 기압, 온도 데이터를 수신한다. 해당 데이터들은 RF 통신 모듈을 통해 지상국에 통신을 통해 실시간 전송된다.

2.3.1 BNO055

BNO055는 가속도계, 지자계, 자이로스코프 데이터를 ARM Cortex-M0 기반의 단일 칩에 집적하고 모든 센서를 fusion 하여 quaternions, Euler angles 혹은 vector에서 작용할 수 있도록 출력한다.

2.3.2 BMP 280

BMP 280은 고도, 기압, 온도 등을 측정할 수 있는 센서이다. 기본적으로 정밀한 측정이 가능하지만 특별한 조치 없이 측정하면 노이즈가 매우 심하여 칼만 필터를 통해 노이즈를 감소시켰다.

2.3.3 RF통신 모듈

이번 발사체의 추진 예상 고도는 400M 이상이다. 그러므로 400M 이상의 고도에서도 정상적으로 통신할 수 있도록 RF 모듈을 선정하였다. 해당 모듈은 ISM Band 대역인 262~264 MHz 대역에서 통신한다. 다른 대역에 비해 최근에 신설된 주파수 대역이기 때문에 혼잡하지 않을 것으로 예상된다. 또한 Line-of-Sight 가 확보된 경우 1000M 이상의 거리 또한 통신이 가능하기 때문에 400M 이상의 고도에서 임무를 수행하는 본 발사체에 유리할 것으로 예상된다.



Fig. 15 BNO055, BMP 280 and RF wireless connection module (from the left)

2.4 통신

로켓 상태를 실시간으로 확인함과 동시에 정보 수집을 위해선 통신 시스템을 구축하는 것이 중요하다. 송·수신 가능 속도와 용량을 고려하여 통신 시스템을 구축하였다. 통신 실험은 개활지를 기준으로 하기 위해 Fig. 17과 같이 교내 운동장에서 진행하였다.

Table 6. Data format.

Name	Data type	Scale
Temperature	integer	°C
Magnetometer	(float,float,float)	micro teslas
Accelerometer	(float,float,float)	m/s ²
Gyroscope	(float,float,float)	rad/sec
Euler angle	(float,float,float)	degree
Gravity	(float,float,float)	m/s ²

2.4.1 통신 프로토콜

자세 데이터는 PAYLOAD 부분에 저장하여 통신한다. 이때 한번에 전송 가능한 용량은 55 Byte이다. 그러므로 자세 데이터를 전송하기 위해서는 데이터를 여러 조각으로 분리하여 여러 번에 걸쳐 송신하고, 수신자는 받은 데이터를 순서대로 조립한다. 또한 데이터 송수신 도중

Packet Loss가 생길 수 있으므로 Timestmap 정보 또한 PAYLOAD에 함께 탑재한다.



Fig. 16 data packet.

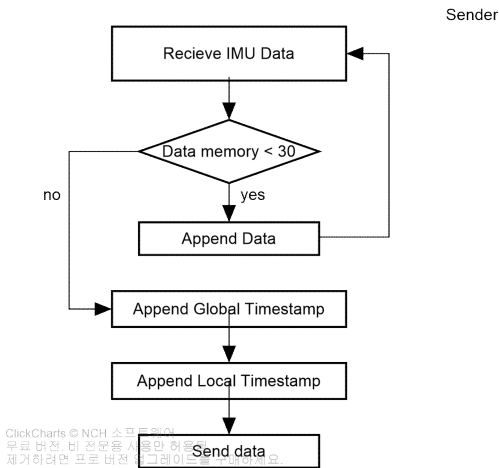


Fig. 17 sender protocol.

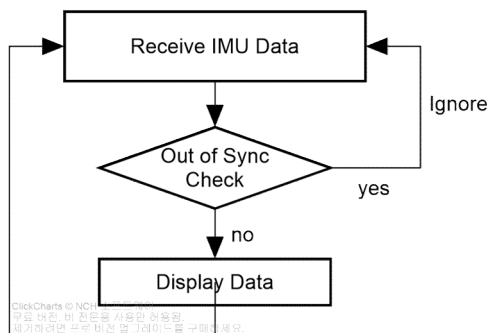


Fig. 18 receiver protocol.

2.4.2 통신 실험 결과

발사 대회장과 유사한 개활지에서 실험을 진행하였다. 200m 거리에서 정상적인 속도로 송수신하는 것을 확인하였다. 추후 목표 고도인 400m 거리에서도 실험을 진행하여 통신의 신뢰성을 높이 예정이다.



Fig. 19 radio communication test location

2.5 모니터링 시스템

앞서 언급한 센서에서 얻은 데이터를 통신으로 전달받고 이를 실시간 처리하여 시각화한 Dashboard를 제작하기로 하였다. Fig. 19은 Figma를 사용한 프로토타입이며 추후 Unity로 제작될 예정이다. 주요 기능은 총 4가지로 구성된다.

- ① BNO055로 받은 자세 데이터를 로켓 모형에 실시간 적용하여 시각화하였다.
- ② 서보를 움직일 수 있는 강제 사출(Emergency ejection) 버튼을 제작하였다.
- ③ UART 통신 기반 RF 통신 시스템을 사용하므로 UART 연결 확인을 위한 UI를 제작하였다.
- ④ BNO055에서 얻은 roll, pitch, yaw 값과 BMP280에서 얻은 온도, 기압, 고도를 시각화하였다.

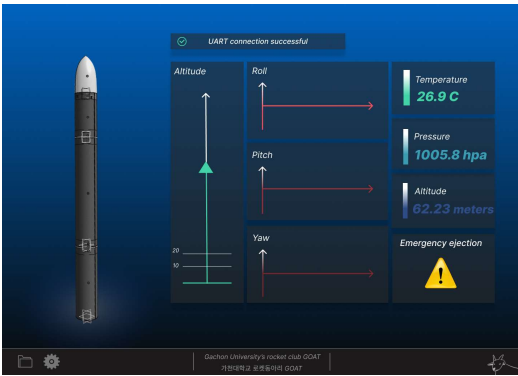


Fig. 20 Dashboard made with figma.

2.5.1 수신데이터 처리

로켓에 탑재된 라즈베리파이로부터 얻은 데이터는 지상국을 거쳐 PC에 전달되어야 한다. 이 과정에서 시리얼 포트가 사용된다.

2.6 자세 제어

핀을 장착한 발사체는 상승 중 roll 축 회전을 한다. 안정적인 통신 데이터와 영상을 수집하기 위해서 발사체의 roll 축 회전을 reaction wheel을 통해 제어한다. 제어에는 PID control을 사용한다.

2.4.1 Reaction wheel

Reaction wheel은 로켓의 roll 축 회전을 제어하기 위한 장치이다. 로켓의 회전을 상쇄시키는 방향으로 회전해 각 가속도를 발생시켜 회전을 제어한다. IMU 센서인 BNO055로부터 받는 자세 제어를 기반으로 회전하지 않도록 작동한다. DC 모터가 Fig. 21과 같이 생긴 wheel을 회전시킨다.



Fig. 21 PID control test equipment.

2.4.2 PID control

Reaction wheel은 PID control을 이용해 발사체의 roll 축 회전각을 0°로 제어한다.

2.5 발사 시퀀스

발사체 발사 시퀀스는 발사 전 로켓의 상태 확인부터 발사체가 낙하산을 사출하고 회수되는 과정까지를 의미한다.

2.5.1 전자 장비 작동 확인

장비 점검 프로그램을 통해 발사 준비 여부를 확인한다. 해당 프로그램은 통신을 통해 센서에서 정상적으로 데이터를 수신하는지 확인한다. 해당 프로그램은 발사체 조립부터 작동하며 문제가 생기면 접촉 상태 및 센서 작동 여부를 재확인한다.

2.5.2 점화

내부 장비의 정상적인 작동이 확인되면, 이그나이터를 이용하여 점화한다. 점화가 이루어지지 않으면 30초 이상 대기 후 이그나이터를 재설치한 후 재점화한다.

2.5.3 비행

비행 중 수집한 데이터를 지상국으로 전송하고, SD카드에 전송 정보를 포함한 영상 데이터를 저장한다. 지상국에서는 실시간으로 발사체에

서 보내오는 신호를 확인 및 분석한다. 영상 데이터는 회수 후 확인한다.

2.5.4 낙하산 사출

낙하산 사출은 다음 3가지 조건을 통해 이루어진다. ① 직전 고도보다 낮은 값이 측정 ② 최고 고도 도달 예상 시간 초과 ③ 지상국에서 강제 사출 신호 전송. 이 중 한 가지를 만족하면 사출되도록 하였다. 1번과 2번이 작동하지 않고 추락하게 된다면, 지상국에서 강제 사출 신호를 전송한다.

2.5.5 착륙 및 회수

발사체가 지상에 내려오면 회수팀은 발사체 상태 확인 후 지상국으로 발사체를 회수한다. 이후 통신으로 전송받은 데이터를 저장 및 분석하고, SD 카드의 데이터 또한 같은 방식으로 처리한다. 수집한 데이터는 향후 발사를 위한 자료로 사용한다.

2.6 비행 시뮬레이터

최소한의 실험으로 목표한 결과를 도출하는 것은 효율적인 연구에 있어서 중요하다. 이에 더하여 로켓은 특수한 경우를 제외하면 발사 실험에 많은 제약이 있다. 특히 자세 제어를 위한 연구는 많은 시행착오가 수반되는 분야이다. 따라서 본 연구팀은 TMS를 통해 수집한 추력 데이터를 활용할 수 있는 비행 시뮬레이터를 개발하였다. 이는 TVC (Thrust Vector Control)를 가상 공간에 구현하는 것으로, yaw, pitch 축 제어를 통해 로켓의 자세 제어를 할 수 있다. 이는 비행 안정성 향상뿐만 아니라 추후 도전할 역추진 제어에 유의미한 사전 정보를 제공할 것으로 기대한다. 센서와 액추에이터의 반응 시간 등을 사전에 수집하여 실제로 구현할 때 필요한 장비의 성능 정보를 수집할 수 있을 것으로 기대한다.

2.7 강화 학습

발사체는 수직 방향으로 추진하며 기상 상황에 영향을 받는다. 또한 직접 제작한 발사체의

경우 해당 발사체의 정보를 사용 시뮬레이터에 적용하기 어렵다는 점이 존재한다. 그러므로 3D 모델링 데이터와 추력 데이터를 활용한 강화학습 TVC 제어 시스템을 제안한다.

강화학습은 환경과 Agent의 상호작용을 통해 시행착오를 겪게 되고, 이 경험을 통한 학습을 하게 된다. 그 결과로 Agent는 주변 환경의 불확실한 요소를 극복하기 위한 방법을 모색한다.

2.7.1 정책

정책은 특정 시점에서 Agent가 취해야 할 행동을 정의한다. 일반적으로 정책은 확률론적으로 행동을 선택할 수 있으므로 강화학습의 핵심적인 요소이다.

2.7.2 보상

Agent가 성취해야 하는 목표이다. 해당 발사체의 경우 기울어짐이 적을수록 더 높은 보상을 얻게 되고, 높이 올라갈수록 더 높은 보상을 얻게 된다.

2.7.3 환경

환경은 Agent 이외의 모든 것을 의미하며, Agent가 제어할 수 없다는 것을 의미한다. 그렇기에 해당 시스템에서 환경은 Agent를 둘러싸고 있는 대기가 된다. 또한, TMS로 얻은 추력 데이터 또한 환경이 된다. 이때, 대기의 경우 일정 시간마다 랜덤한 방향으로, 랜덤한 힘을 가하여 대기 상태를 모사하였다.

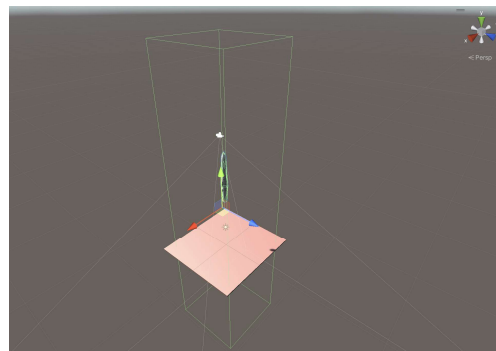


Fig. 22 rocket object in unity.

2.7.4 행동

발사체는 수직 방향으로 추진하여 TVC를 하게 된다. 이때, 노즐의 회전각은 pitch 축으로 -10

~ 10 yaw 축으로 $-10 \sim 10$ 회전할 수 있으므로 $(-10 \sim 10, -10 \sim 10)$ 의 연속적인 2차원 값이 Agent의 행동이 된다.

2.8 Proximal Policy Optimization Algorithm

PPO 알고리즘은 Model-Free based learning의 한 종류이다. 이전 알고리즘의 경우 최적화 진행 후 사용된 데이터는 버려지게 되고, 이는 정책에 대해 높은 분산으로 인한 학습 성능 저하가 생기게 된다. PPO는 이전에 사용한 경험을 재활용하기 때문에 학습 측면에 있어 효율성을 가져오게 된다.

Table 7. Hyperparameter.

Hyper parameter	Value
batch_size	64
Buffer_size	2048
learning_rate	0.0003
beta	0.001
epsilon	0.2
lambd	0.99
num_epoch	3
learning_rate_schedule	linear
beta_schedule	linear
normalize	true
hidden_units	128
num_layers	2
vis_encode_type	simple
memory	null
goal_conditioning_type	hyper
deterministic	false

2.8.1 ML-Agent

유니티 ML-Agent는 게임을 포함한 다양한 시뮬레이션에서 활용하기 위한 지능형 Agent를 훈련 시키는 환경을 제공하는 툴이다. 해당 툴을 활용하여 환경에 따른 Agent를 학습시켰다.

2.8.2 시뮬레이터 결과

Cumulative Reward의 경우 Agent가 획득한 Reward의 누적 합이다. 시간에 따라 점점 높아지는 경향을 갖고 있는 것을 보아 학습이 맞게 되고 있는 것으로 추정할 수 있다. 또한 Value Lose의 경우 시간에 따라 낮아지는 것을 볼 수 있고, 이 또한 학습이 잘 진행되었다는 것을 볼 수 있다.

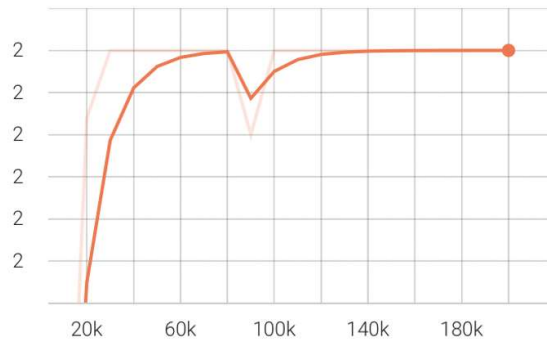


Fig. 23 cumulative Reward

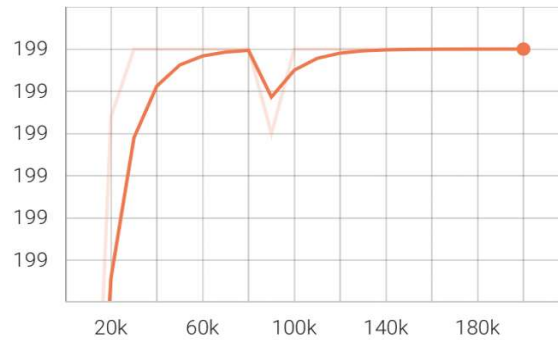


Fig. 24 episode length.

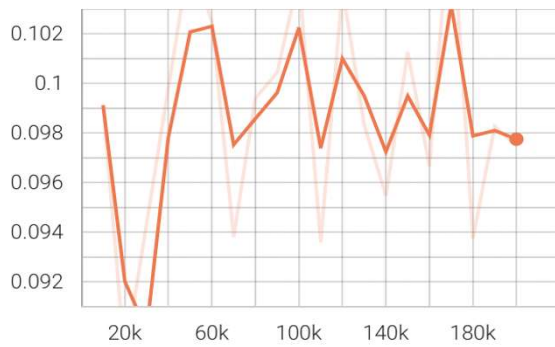


Fig. 25 policy loss.

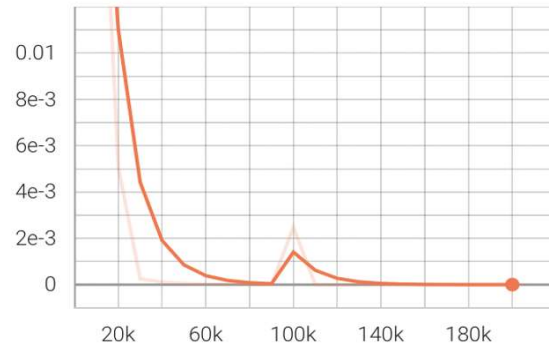


Fig. 26 value loss.

3. 결 론

이번 연구를 통해 기존보다 발전한 엔진을 개발할 수 있었다. 이전보다 더 높은 신뢰성을 가진 실험 값을 얻었고 이를 통해 향후 개발할 다양한 분야에 적용할 수 있었다. 모니터링 시스템은 보다 체계화된 발사 시퀀스와 더불어 정보

수집뿐만 아니라, 실제 발사 현장에서 유용한 도구로 사용할 수 있을 것으로 기대된다. 이에 더하여, 인공지능 강화학습을 통한 제어 시스템은 시뮬레이터뿐만 아니라 실제 제어 시스템으로 사용할 수 있는 가능성을 보여주었다. 해당 분야는 지속적으로 발전시켜 PID 제어에 사전 정보를 얻는 것을 넘어 실제 비행에도 적용할 예정이다. 더 나아가 많은 사람들이 사용할 수 있도록 오픈 소스로 공개해 연구자들이 쉽게 사용할 수 있도록 할 계획이다.

후 기

논문 작성에 도움을 주신 모든 분들께 감사드립니다.

References

1. I Hossam, Sh Saleh and Kamel, "Review of challenges of the design of rocket motor case structures", IOP Conference Series: Materials Science and Engineering 18th, pp. 7, 2019.
2. Sutton, G.P., *Rocket Propulsion Elements*, 6th ed., John Wiley & Sons Inc., New York, N.Y., U.S.A., 1992.
3. Proposal For The Inclusion of KNO₃/SUGAR PROPELANTS IN THE TRA EXPERIMENTAL ROCKETPY PROGRAM, Stuart Lesile & James Yawn, 2002.
4. The Descriptive Geometry of Nose Cones, Crowell Sr., Gray A., 1996, pp.9.
5. Juliani, A., Berges, V., Teng, E., Cohen, A., Harper, J., Elion, C., Goy, C., Gao, Y., Henry, H., Mattar, M., Lange, D. (2020). Unity: A General Platform for Intelligent Agents. arXiv preprint arXiv:1809.02627. <https://github.com/Unity-Technologies/ml-agents>.
6. Sutton, Richard S., and Andrew G. Barto. Reinforcement learning: An introduction. MIT press, 2018.
7. Schulman, John, et al. "Proximal policy optimization algorithms." arXiv preprint arXiv:1707.06347 (2017).