

Virtual-Constructive 시뮬레이션 연동을 활용한 공중전 전투 실험

김동준 · 신용진 · 안경수 · 김영곤 · 문일철 · 배장원[†]

Virtual-Constructive Simulation Interoperation for Aircombat Battle Experiment

Dongjun Kim · Yongjin Shin · Kyeong-Soo An · Young-Gon Kim · Il-Chul Moon · Jang Won Bae[†]

ABSTRACT

Simulations enable virtually experiencing rare events as well as analytically analyzing such events. Defense modeling and simulation research and develops the virtual and the constructive simulations to support these utilizations. These virtual and constructive(VC) simulations can interoperate to simultaneously virtual combat experience as well as evaluations on tactics and intelligence of combat entities. Moreover, recently, for artificial intelligence researches, it is necessary to retrieve human behavior data to proceed the imitation learning and the inverse reinforcement learning. The presented work illustrates a case study of VC interoperations in the aircombat scenario, and the work analyze the collected human behavior data from the VC interoperations. Through this case study, we discuss how to build the VC simulation in the aircombat area and how to utilize the collected human behavior data.

Key words : Artificial Intelligence, Dog fighting, Air Warfare, Reinforcement Learning, VC Interoperation, Virtual Simulation, Constructive Simulation

요 약

시뮬레이션을 통하여 경험하기 어려운 사건을 가상적으로 체험할 수도 있으며, 분석을 수행할 수도 있다. 국방시뮬레이션 분야에서 이런 체험을 기반으로 Virtual 시뮬레이션이 연구 개발되고 있으며, 분석을 위하여 Constructive 시뮬레이션이 연구 개발되고 있다. 이런 시뮬레이션을 연동하여, VC(Virtual-Constructive) 시뮬레이션 환경을 구성할 수 있고, VC 시뮬레이션 환경은 전투 환경의 가상적인 경험과 교전개체의 지능적인 전술을 동시에 시험할 수 있는 환경이다. 또한, 최근의 인공지능 연구를 위해, 사람의 행동을 학습하는 이미테이션 학습 혹은 역강화학습 분야는 VC 환경에서 수집된 인간 행동 데이터를 필요로 한다. 제시된 연구는 공중전 분야에 VC 시뮬레이션 환경의 사례를 보여주며, 이를 통해 수집된 인간 행동 데이터의 특징을 분석하고 있다. 본 논문을 통하여, 공중전 분야 VC 시뮬레이션 환경이 어떻게 구축될 수 있으며, 인공지능 학습을 위하여 어떻게 활용될 수 있는지 보여준다.

주요어 : 인공지능, 공중전, 공중 무기체계, 강화학습, VC 연동, Virtual 시뮬레이션, Constructive 시뮬레이션

1. 서론

시뮬레이션은 가상세계를 제공하여, 현실에서 발생시키기 어려운 시나리오를 개연성 있게 실현할 수 있는 장점이 있다. 이런 장점을 통하여, 사람들이 새로운 경험을 수행할 수도 있으며, Rare event(즉, 매우 드물게 일어나는 사건)를 통계적으로 분석할 수 있는 가상 데이터도 제공한다. 특히, Rare event의 대표적인 사례를 활용하는 분야는 군사 분야이다. 전쟁은 쉽게 체험하기 어렵고, 발

* 이 논문은 2020년도 한화시스템(주)의 재원을 지원 받아 수행된 연구임

Received: 7 September 2020, Revised: 12 March 2021,
Accepted: 13 March 2021

[†] Corresponding Author: Jang Won Bae

E-mail: jangwon_bae@koreatech.ac.kr

School of Industrial Management,

Korea University of Technology and Education

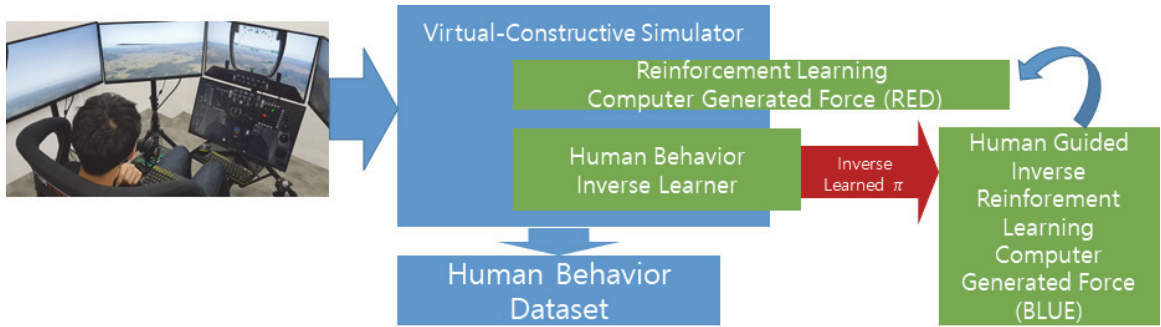


Fig. 1. Utilization of human behavior dataset from virtual-constructive simulation for inversely learned computer generated force

생하는 빈도도 매우 낮다. 그러나 그 결과의 중요성 때문에, 전쟁 수행 인원에게 가상의 체험을 제공하여 훈련 시켜서 현실에서 Rare event를 발생했을 때, 효율적으로 대처할 수 있게 대비시킨다. 현실과 유사한 가상의 전쟁을 무수히 반복하여, 중요한 분석결과물을 유의미하게 생성한다.

이렇게 국방 시뮬레이션 분야에서 활용되는 사례는 사용자의 경험을 위한 Virtual 모델링 및 시뮬레이션과 분석자의 데이터 습득을 위한 Constructive 모델링 및 시뮬레이션으로 나누어 볼 수 있다. Virtual 모델과 Constructive 모델은 활용의 주체, 용도가 다르기 때문에 개발 시 강조되는 기술적 난점 및 이론적 배경도 상이하다. Virtual 모델의 경우, 사용자가 가상 체험을 현실적으로 받아들일 수 있는 인터페이스, 가상 환경, 시뮬레이션 내 가상 개체의 행동 현실성이 중요하다. 이를 기반으로 이론적 배경은 3D 모델링, 인간-컴퓨터 상호작용(human computer interaction), 지능적 가상 행동을 일으키는 에이전트 모델링 등이 존재한다. 이외는 상이하게 Constructive 모델은 분석자가 Rare event를 개연성 있게 다수 생성하는 과정이 필요하다. 그러므로 대규모 시뮬레이션, 병렬 시뮬레이션, 개연성 있는 시뮬레이션 진행을 위한 에이전트 행동 모델링 등이 필요하다. 또한 모의되는 사건을 체계적으로 표현하기 위한 형식론 및 시스템 공학적 이론 배경이 필요하다.

지금까지 Virtual 및 Constructive 모델은 주로 개별 활용되는 사례가 많았다. VC(Virtual-Constructive) 연동(interoperation)을 통한 훈련 모델을 개발하여, 현실에 활용한 사례들도 다수 존재한다. 이런 훈련 목적의 VC연동은 시뮬레이션내의 사용자 행동 데이터를 수집 보다, 지정된 시나리오를 체계적으로 수행하며, 사용자가 훈련과정을 가상적으로 체험하는 것을 주된 목적으로 삼았다.

최근 인공지능의 발달은 전투개체가 사람과 같은 행동을 보일 수 있다는 가능성을 보여주었다. 특히 역강화학습(inverse reinforcement learning, Abbeel 등, 2004) 및 모방학습(imitation learning, Ho 등, 2016)의 분야는 사람의 행위 데이터를 활용하여, 인공지능이 주어진 상황에 맞게 사람 같은 행위를 보이는 연구를 제시하였다. 이러한 연구에는 다양한 상황에서의 사람의 행동 및 상황 인식 데이터가 필수적이지만, 군사 분야와 같은 Rare event를 고려하는 분야에서는 데이터 확보가 현실적으로 불가능하다. 이러한 관점에서, 본 논문은 VC 연동이 가상 상황에서의 사용자의 행동과 관련된 데이터를 수집할 수 있는 도구로 활용될 수 있는 가능성을 제시한다. Fig. 1은 위에서 설명한 VC 연동을 통해 획득한 데이터를 인공지능의 학습과정에 어떻게 활용할 수 있는지 설명한다.

논문의 시뮬레이션 시나리오는 1대1 기총기반 공중전 상황(Bonanni, 1993; Shaw, 1985)으로, 공중 기체의 의사6자유도기반의 기동 모델 및 살상 원뿔 형태의 기총 모델의 상세함을 제공한다. 사용자 인터페이스는 기체의 조종석과 유사하게 설계되어, 실제 상황의 입력을 재현할 수 있도록 설계되고, 사용자는 이를 이용하여 적기를 추적하는 임무를 수행한다. 이러한 과정에서 사용자와 적기의 상태정보 및 사용자의 제어 입력 정보를 모두 수집할 수 있고, 본 논문은 수집된 데이터에 대한 분석 결과를 제공한다. 또한, 시뮬레이션을 통해 수집된 데이터를 역강화학습의 학습 데이터로 활용할 수 있는 방법에 대해서 설명한다.

2. 기존 연구

공중전 교전은 다양한 시뮬레이션을 통해 연구가 수행되었는데, 한승엽 등(2016)에 따르면, Score 기반의 교전

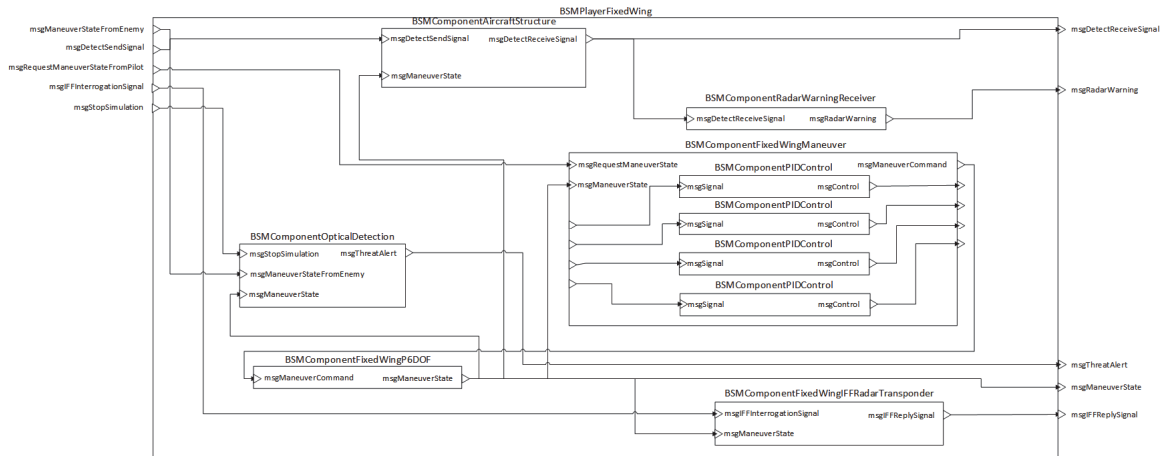


Fig. 2. DEVS Coupled Model of Aircraft

을 Pay-off 행렬로 모델링할 수 있었다. 이런 관점에서 사람과 인공지능의 상호작용을 게임이론으로 표현할 수 있으며, 이를 VC연동을 통해 데이터 수집할 수 있다.

기존 공중전 교전을 강화학습으로 수행한 사례는 이동진 등(2009)에 의해서 소개된 경우가 있다. 이때의 강화학습의 방식은 Markov Decision Process(Sutton 등 1998) 기반의 강화학습이다. 제시된 역강화학습 종류 중의 하나는 Inverse Markov Decision Process가 존재하여, 기존의 강화학습 기반 사례의 역방향 학습이 수행된다는 점을 알 수 있다. 위의 사례에서 사람은 개입의 여지가 없으나, 본 논문에서 수집되는 VC연동 사용자 데이터를 통해, 역강화학습을 수행할 수 있다.

해외의 경우 공중전에 인공지능을 부여한 사례들이 더욱 많이 존재하지만, 역강화학습을 제시한 경우는 매우 드물다. 예를 들어, Ernest 등(2016)의 사례도 Fuzzy 로직 기반의 인공지능을 활용하여, 사람을 대체하는 연구를 한 것이지, 사람의 행동을 배우려는 시도는 아니었다. 최근의 깊은 강화 학습(deep reinforcement learning) 기반의 교전 사례 역시, 사람의 행동 데이터를 활용하는 경우는 없었다. 예를 들어, Toghiani-Rizi 등(2017)에 따르면 컴퓨터 생성군(computer generated force)이 인공지능으로 통제되어 지상 작전을 수행하는 사례가 발표되었지만, 이는 Constructive 시뮬레이션을 통해 반복 학습된 것이지, 사람의 행동데이터에 기반을 두지는 않았다.

3. Virtual-Constructive 연동 시뮬레이션

위에서 논의한 바와 같이 본 논문의 시뮬레이션은 VC

연동 기반의 공중전 시뮬레이션이다. 본 장은 시뮬레이션의 Virtual 모델, Constructive 모델 및 VC 연동 부분의 모델을 설명한다.

3.1 Constructive 모델

Fig. 2는 시뮬레이션으로 구현된 Constructive 모델의 구조를 표현하고 있다. Fig. 2에서 제안된 모델은 1대1 교전에 활용된 하나의 기체를 보여주는 것이며, Fig. 2의 모델에서 생성되는 입력과 출력을 교차 연결하여 1대1 교전 DEVS Coupled 모델이 구성된다. 그리고 교차 연결된 DEVS Coupled 모델이 Outmost 모델이 된다.

제안된 시뮬레이션은 1대1 공중전 시나리오를 기반으로 하고 있다. 즉, 아군 기체(blue)는 적 기체를 추적하는 역할을 수행하며, 적 기체(red)는 사용자 기체를 회피하는 시나리오이다. 이 2기의 교전 개체(즉, blue와 red 조종사)는 모두 같은 모델 구조를 가지고 있다. 기존 국방 시뮬레이션의 교전개체 모델링과 같이 조종사 모델은 기동, 탐지, 교전, 피해평가를 수행할 수 있는 세부 모델을 포함하고 있다. 위와 같은 교전, 개체, 세부동작의 계층적 구조를 표현하기 위하여 DEVS 형식론을 활용하였으며(Ziegler 등, 2000), 이런 계층적 구조는 조종사 모델을 구성하는데 활용되었다. Fig. 3은 조종사 모델(Aircraft Pilot)의 구조를 표현하고 있다. 특히, 제시된 VC연동에서는 기동 교전만 활용하였으나, 구현된 전체 모델은 미사일, 채프, 플레어의 투사도 모델링 되어 있다. 세부 절에서 조종사 모델의 세부 행동(의사 결정, 탐지, 기동, 교전, 피해평가)에 대한 모델링 내용을 설명한다.

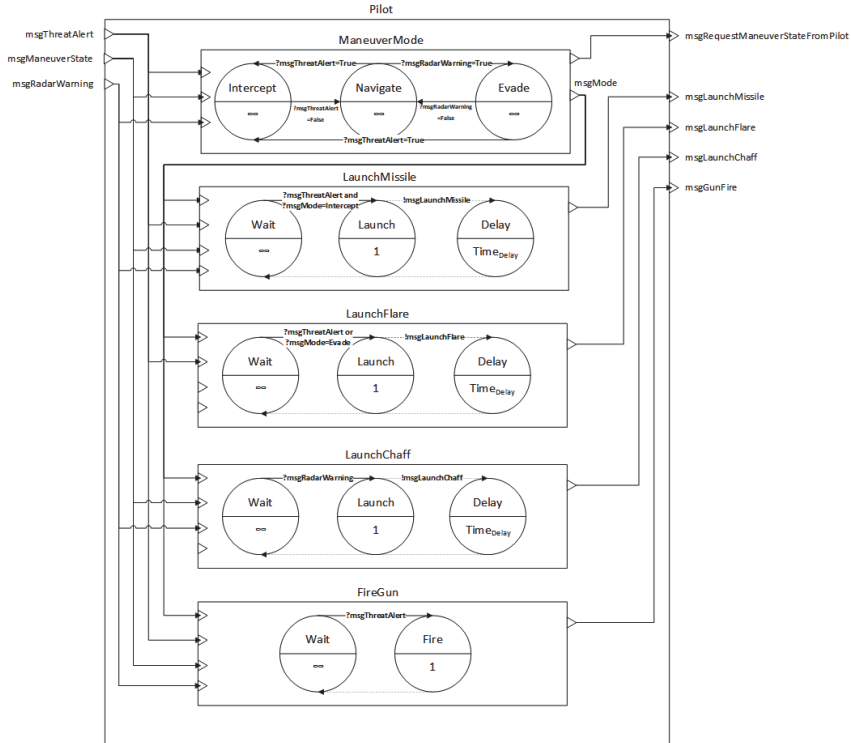


Fig. 3. DEVS Coupled and Atomic Model of Aircraft Pilot Model

3.1.1 의사 결정 모델

Fig. 3은 조종사 모델의 개략적인 의사결정 구조를 표현하고 있다. 조종사 모델의 구체적 의사결정은 DEVS Atomic 모델로 구현되었으며, 이 모델은 Atomic 모델 다이어그램 혹은 시퀀스 다이어그램으로 표현될 수 있다. 또한, 이의 실질적인 구현은 프로그램으로 이루어지는데, 이 프로그램의 상태 천이를 순서도로 표현하면 Fig. 4와 같다. 구체적으로, 아군이 적군을 발견하면 우선 미사일인지 적기체인지 판단하고, 필요시에 회피 기동을 수행한다. 그 후, 본체의 무기 장비(미사일, 플레어, 채프 등)의 상태를 확인한 후, 준비된 무기를 적에게 발사하여 격추를 노린다.

3.1.2 탐지, 피해 평가 및 교전 모델

탐지 모델은 피아 구분 및 상대의 NED(North, East, Down) 좌표에 대한 위치, Roll, Pitch, Yaw에 대한 자세 각 정보를 수집한다. 피해 평가 모델 및 교전 모델은 아군기의 기체 NED좌표계에서 N방향으로 원뿔을 발생시켜, 원뿔이 피해반경에 되도록 설계하였다. 만약 적기가 일정 시간 이상, 피해반경 원뿔에 위치한다면, 시뮬레이션이 끝나도록 설계되었다.

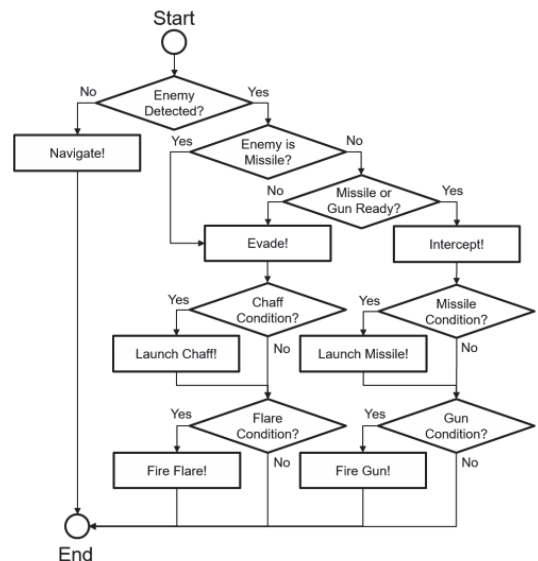


Fig. 4. Flowchart on pilot decision-making by external events. This is implemented in the Pilot's maneuver model by sending "mode" output of decisions to other models, such as gun, missile, etc.

3.1.3 기동 모델

제시된 1대1 교전의 기동 모델은 의사 6자유도 고정의 운동방정식을 활용하였다. 표 1은 운동방정식의 각 변수를 정리한다.

의사 6자유도를 활용한 기동의 제어는 기존 연구논문(문일철 등, 2020)을 따라서 구성하였다. 기본적으로 추력 T_t , 양력 L , 롤 각속도 $\dot{\phi}$ 의 제어를 수행하였다. 추력 T_t , 양력 L , 롤 각속도 $\dot{\phi}$ 의 제어는 Virtual 시뮬레이션의 경우 사용자에게서 추력(thrust, T), 하중계수(load factor, N), 경사각(bank angle, P)을 입력받아 기동모델에 메시지로 전달하였다. 제어에 활용된 L과 $\dot{\phi}$ 는 각각 N과 P에 민감도 수준의 사용자의 설정에 따른 계수를 곱한 결과물이다. T 역시 민감도를 고려할 수 있도록 사용자의 설정에 따른 계수를 곱하였다.

$$\begin{aligned}\dot{V}_t &= (T_t - D)/m - g \sin \gamma_t \\ \dot{\gamma}_t &= \frac{(L + T_t \sin \alpha) \cos \phi - W \cos \gamma_t}{m_t V_t} \\ \dot{\psi}_t &= \frac{(L + T_t \sin \alpha) \sin \phi}{m_t V_t \cos \gamma} \\ \dot{X} &= V_t \cos \gamma_t \cos \psi_t \\ \dot{Y} &= V_t \cos \gamma_t \sin \psi_t \\ \dot{h} &= V_t \sin \gamma_t\end{aligned}$$

또한, Constructive 시뮬레이션의 경우도 마찬가지로 T, N, P를 제어를 생성하여, T_t , L , $\dot{\phi}$ 를 결정한다. 여기서, T, N, P의 생성과정은 회피 알고리즘에서 결정한다.

3.1.4 적 기체의 회피 알고리즘

제안된 시뮬레이션에서 사용자 기체는 적 기체를 추적하는 역할을 수행하며, 적 기체는 사용자 기체를 회피하는 과정을 시뮬레이션 시나리오로 설정하였다. 그러므로 본 논문에서는 아군 기체(공격기)를 실험자가 조종하여

Table 1. Variables of Pseudo 6DOF

No.	Notation	Variables
1	V_t	동체 축에서의 속도 벡터 [m/s]
2	T_t	동체 축에서의 추력 벡터 [N]
3	α	받음각 [rad]
4	γ_t	상하각 [rad]
5	ψ_t	방위각 [rad]
6	X	관성 좌표계에 대한 X축 위치 [m]
7	Y	관성 좌표계에 대한 Y축 위치 [m]
8	h	관성 좌표계에 대한 고도 [m]
9	L	양력 벡터 [N]
10	D	항력 벡터 [N]
11	W	중력 벡터 [N]
12	m_t	동체 질량 [kg]
13	g	중력 가속도 [m/s^2]
14	ϕ	롤 자세각 [rad]

적군 기체(방어기)를 격추시키는 실험을 진행한다.

방어기의 회피 기동은 시뮬레이터에 미리 입력한 알고리즘에 의해 결정되는데, 크게 두 가지 목표를 이루기 위해 알고리즘을 디자인하였다. 첫 번째 목표는 공격기의 진행 방향과 수직 방향으로 방어기가 회피하는 것이다. 두 번째 목표는 방어기의 진행 방향을 공격기가 예측하지 못하도록 하는 것이다. Fig. 5(a)는 방어기의 회피 기동을 도식적으로 설명하고 있다. 방어기는 공격기의 진행 방향과 수직으로 회피하여 결국 공격기를 마주하는 방향으로 움직인다.

목표를 자세히 설명하기에 앞서, 기체의 기동이 어떻게 이루어지는지 설명하도록 하겠다. 방어기는 알고리즘이 찾아 준 목표 Pitch와 목표 Yaw를 달성하기 위해 PID Controller를 사용한다. 먼저, 목표 Pitch를 달성하기 위해 PID Controller의 입력으로 현재 Pitch 값과 목표 Pitch 값을 입력하면 PID Controller는 목표 Pitch를 달성

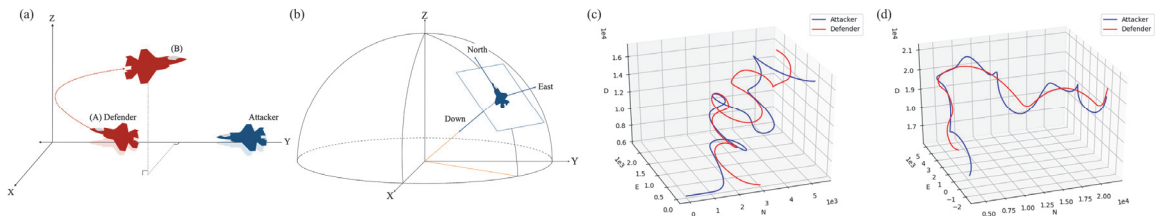


Fig. 5. Evasion rule and simulation result (a) defender evasion rule (b) attacked NED coordinate (c) defender maneuver under 20,000(m) altitude (d) defender maneuver over 20,000(m) altitude where the altitude limit is applied

하기 위한 N 값을 결과로 내보낸다. 또한, 현재 Yaw 값과 목표 Yaw 값을 입력하면 PID Controller는 목표 Yaw를 달성하기 위하여 P 값을 결과로 내보낸다. 마지막으로, PID Controller는 목표 속력에 도달하기 위해 현재 속력과 목표 속력을 입력으로 받아 T 값을 결과로 내보낸다. PID Controller의 출력인 TNP란, 추력인 T , 하중 계수인 N , 경사각인 P 를 의미한다. 주어진 TNP 값을 입력으로 하여, 다음 상태 변수를 의사 6자유도 기동 모델로 얻는다. 즉, PID Controller와 의사 6자유도 모델을 이용하여 목표 Pitch와 목표 Yaw로 근접해가는 기동을 얻을 수 있다.

알고리즘 디자인을 설명하기 위하여, 격추 조건을 살펴보도록 하겠다. 본 실험에서, 아군 공격기의 좌표계를 기준으로 적군 방어기가 일정 사거리와 일정 자세각 안에 위치하는 경우, 적군 방어기가 격추되었다고 간주한다. 회피 기동은 격추 조건을 만족시키지 않도록 움직여야 하기 때문에, 아군 공격기로부터의 거리 및 자세각을 일정 기준 이상으로 만드는 것을 목표로 한다. 따라서 방어기가 격추를 피하기 위한 최적 회피 기동은, 첫 번째로 자세각을 늘리기 위하여 공격기의 이동 방향과 직각이 되는 방향으로 회피하는 것이다.

Fig. 5(b)는 공격기를 기준으로 한 NED 좌표계를 나타낸다. NED 좌표계란, 각각 진행 방향을 나타내는 N (north) 축과 진행 방향으로부터 수직으로 오른쪽을 나타내는 E (east) 축, 그리고 진행 방향으로부터 수직으로 아래를 나타내는 D (down) 축을 의미한다. 방어기는 방향 선정을 위해 공격기를 기준으로 한 NED 좌표계에서 E 축과 D 축으로 직각이 되는 방향을 목표 Pitch와 목표 Yaw로 잡는다. 목표 Pitch를 통해 D 축 방향으로 공격기로부터 수직으로 이동하기 위한 방어기의 방향이 결정되며, 목표 Yaw를 통해 E 축 방향으로 공격기로부터 수직으로 이동하기 위한 방어기의 방향이 결정된다. 방어기의 목표 Pitch와 목표 Yaw를 공격기의 진행 방향의 직각으로 설정함으로써, 방어기는 공격기로부터의 자세각을 최대화한다.

두 번째 알고리즘의 목표는 방어기의 진행을 공격기가 예측하지 못하게 하는 것이다. 공격기가 방어기의 회피 경로를 쉽게 예측하면, 예측되는 미래 회피 경로로 방향을 선회하여 현실적이지 않은 실험을 할 가능성이 높기 때문에, 알고리즘에 의한 회피 경로를 예측 불가능하도록 설정하였다. 알고리즘은 예측 불가능한 회피 경로를 위해 목표 Pitch와 목표 Yaw의 방향을 시간에 따라 전환시킬 수 있다. 본 논문에서는 목표 Pitch는 공격기의 NED 좌표계를 기준으로 위 방향($-D$ 방향)으로 고정시키고 목표

Yaw의 방향을 주기성을 가지고 -90° 및 90° 로 전환시킨다. 즉, 방어기는 시간에 따라 왼쪽으로 이동하다가 일정 시간이 되면 오른쪽으로 이동하는데, 변경되는 주기는 5초에서 15초 사이에서 랜덤하게 결정하였다. 변경 주기 범위인 5초와 15초는 실제 실험을 하였을 때 공격기가 방어기의 움직임을 예측하는 시간이 짧게는 5초, 길게는 15초 정도 걸리기 때문에 VC 시뮬레이터의 상황을 고려한 범위로 설정하였다. 최종적으로 방어기는 Fig. 5(c)의 붉은 궤도와 같이 일정 시간 동안 나선형을 그리며 좌상단 혹은 우상단으로 회피 기동을 하게 된다.

방어기는 예외의 경우를 제외하고는 위에서 설명한 바와 같이 회피 기동한다. 현실을 반영하기 위하여 우리는 회피 기동에 몇 가지 예외를 두었다. 첫째, 현실에서는 기체의 안정적인 기동을 위한 최대 고도가 존재한다. 본 시뮬레이터에서는 방어기가 상승할 수 있는 최고 고도를 2만 미터로 두어, 최고 고도에 도달하면 상승하는 회피 기동을 멈추고 단순 하강 기동을 하도록 설정하였다. 고도가 떨어져서 다시 2만 미터 이내로 들어오면 회피 기동을 하도록 디자인하였다. Fig. 5(d)에서 이러한 최대 고도에서의 회피 기동을 볼 수 있다. 둘째로, 기체가 지면에 부딪히는 경우를 고려하여, 고도가 2천 미터 이하로 떨어질 경우, 위에서 설명한 회피 기동을 멈추고 단순 상승 기동을 하도록 디자인하였다.

3.2 Virtual 모델

제안된 시뮬레이션은 1대1 공중전 시나리오를 기반으로 하고 있으며, Virtual 상황의 시현은 X-Plane의 버전 11을 활용하였다. Constructive 모델과 Virtual 모델의 통신은 TCP/IP의 커스텀 프로토콜을 활용하였으며, 이는 X-Plane의 연결 프로토콜을 기본적으로 차용한다. 아직까지 제시된 VC연동의 시뮬레이션 기반 연동체계(예를 들어, DIS, HLA/RTI 등)는 고려되지 않았다. 그러나 지금의 TCP/IP수준으로도 Dead Reckoning현상은 발생하지 않았으며, 1대1 기총기반 공중 교전을 정상적으로 수행하는데 어려움은 없었다. 연동 차원의 연구는 기존에 다수 존재하며, 본 논문은 VC 연동을 통하여 인공지능 학습 데이터를 생성하는데 주목한다. 이런 관점에서, Virtual 모델을 활용한 사용자의 행동 규칙을 제시해야 한다. 또한 기본적인 Virtual 모델의 인터페이스 설정을 제시한다.

3.2.1 아군 기체의 추적 행동

본 실험은 현실에서는 주어져 있지 않은 보상 함수를 데이터로부터 추론하여, 최적의 추적 경로를 찾는 것을



Fig. 6. Virtual simulation capture (a) Lag pursuit identified by the flight cloud (b) pure pursuit (c) lead pursuit

목표한다. 공격기가 방어기의 후방에 위치한 경우, 정성적으로 분석할 수 있는 추적 모드는 Fig. 6(a)에 나온 바와 같이, 지연 추적(lag pursuit), 일치 추적(pure pursuit), 선도 추적(lead pursuit)으로 구성된다.

지연 추적은 방어기의 궤도의 후미를 따라가기 때문에 세 가지 추적 방법 중 가장 방어기의 궤도를 충실하게 따라가는 방법이다. 이러한 지연 추적은 안정적인 추적 방법으로, 적기의 속도보다 공격기의 속도가 빠를 경우, 공격기가 방어기의 궤도를 벗어나 방어기를 놓치지 않고 안정적으로 추적할 수 있다. 반면에, 지연 추적은 방어기가 성공적인 회피 기동을 할 경우, 일반적으로 방어기를 격추시키지 못하기 때문에 추적을 위해서 일치 추적이나 선도 추적을 사용한다.

일치 추적은 공격기가 적기의 현재 위치를 추적하는 방법으로써, 공격기와 방어기의 자세각을 격추 조건 이내로 만들기 위해 일치 추적을 이용한다. 또한, 일치 추적은 공격기와 방어기의 거리를 좁히는 데 지연 추적보다 유리한데, 그 이유로는 방어기가 선회 회피를 할 경우, 공격기의 선회 반경이 지연 추적보다 줄어들어 빠르게 적기와의 거리를 좁힐 수 있기 때문이다. 그러나 공격기가 방

어기보다 선회 성능이 떨어질 경우, 충분한 선회가 이루어지지 않아 적기의 선회 궤적을 가로질러 방어기를 놓칠 가능성이 존재한다.

선도 추적은 공격기가 적기의 미래 위치를 예측하여 추적하는 방법으로, 예측이 성공적인 경우 가장 좋은 추적 방법이라고 할 수 있다. 하지만, 예측에 실패할 경우, 다른 추적 방법보다 방어기를 놓칠 가능성이 가장 큰 추적 방법이기도 하다. 선도 추적은 성공적일 경우, 일치 추적보다 더 선회 반경을 줄여 멀리 떨어진 적기를 일치 추적보다 더 빠르게 쫓아갈 수 있다. 공격기는 선도 추적하기 위해 일치 추적에서보다 더 좋은 선회 능력이 필요하며, 그렇지 않은 경우 방어기의 선회 궤적을 가로지를 가능성이 가장 큰 추적 방법이다.

Fig. 7은 시뮬레이션 환경 아래 실험자(공격기, blue fighter)가 방어기(red fighter)를 쫓는 과정에서 지연 추적, 일치 추적 그리고 선도 추적을 전략적으로 선택하는 모습을 보여준다. Fig. 7(b)와 (c)는 같은 경로를 다른 각도에서 시각화한 것으로, 실험 초기에는 실험자가 지연 추적을 통해 거의 동일한 경로로 방어기의 뒤를 쫓는 것을 확인할 수 있다.

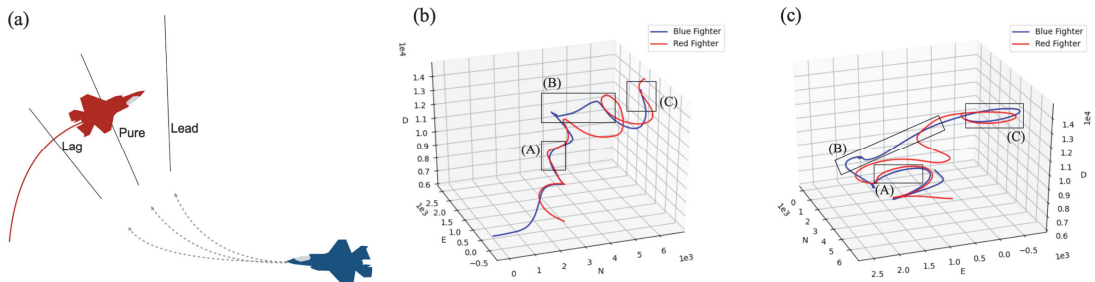


Fig. 7. Pursuit rule and simulation result (a) Basic fighter maneuver pursuit (b) simulation result without 180 degree evasive turn by defender (c) simulation result with 180 degree evasive turn by defender

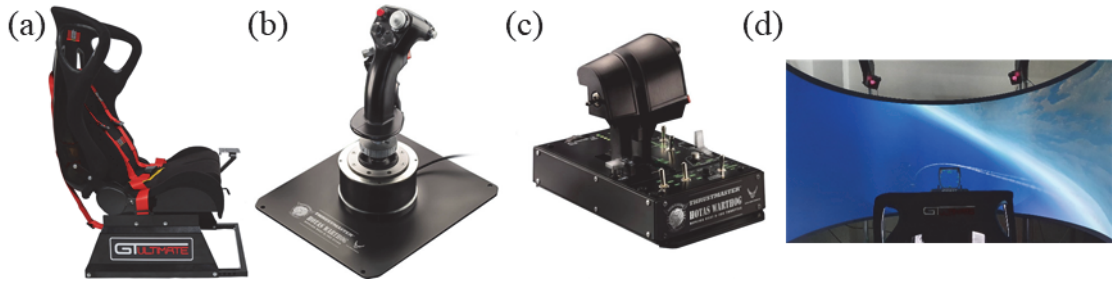


Fig. 8. Virtual simulator hardware (a) Next Level Racing GTULTIMATE SEAT ADD ON chair (b) Thrustmaster Hotas Warthog Flight Stick (c) Thrustmaster Hotas Warthog Dual Throttles (d) SIMPIT Technology Centurion PRO 270

실험자는 총 두 지점(A, C)에서 격추 시도를 하였으나, 첫 번째 시도에서는 격추에 실패하였다. 먼저, A 지점에서 실험자는 일치 추적을 통해 방어기의 선회 궤도 안으로 진입하여 방어기와와 사거리를 좁혔다. 그러나 실험자는 방어기를 격추할 수 있는 사거리에 진입하였음에도 불구하고, 여전히 높은 속도로 인해 격추 조건을 만족하는 자세각을 만들지 못하였다. 따라서 첫 번째 격추 시도는 실패하였고, 실험자는 A 지점과 B 지점 사이에서 방어기의 궤적을 잠시 놓치게 되었다. 실험자는 B 지점에서는 방어기와 거리가 벌어지자 방어기의 거시적인 회피 기동을 예측한 후, 선도 추적 궤적으로 기체를 최고 속도로 올려 방어기를 빠르게 따라잡았다. 실험자는 방어기에 근접한 후, 첫 번째 격추 시도와 달리 감속을 하고 지연 추적으로 방어기를 뒤쫓다가, C 지점에서 일치 추적으로 두 번째 격추 시도를 하여 격추에 성공하였다.

3.2.2 Virtual 모델의 인터페이스

Virtual 시뮬레이터는 Fig. 7(b) 및 (c)로 입력받은 사용자 입력값(TNP)을 Constructive 시뮬레이터로 전달하고, Constructive 시뮬레이터는 받은 입력에 대한 다음 아군 기체의 위치와 적군 기체의 위치를 계산하여 Virtual 시뮬레이터로 전달한다. 위치를 전달받은 Virtual 시뮬레이터는 위치를 시각화하여, 사용자로부터 다음 TNP값을 입력받는다. 이를 위해, 본 연구에서는 Virtual 시뮬레이터를 두 가지 요소로 구성하였다.

첫째, Virtual 시뮬레이터는 Constructive 시뮬레이터로부터 받은 피아 기체의 위치를 시각화해주는 공개 소프트웨어인 Xplane을 사용하였다. Xplane은 그 자체로 비행기 조종 Constructive 시뮬레이터이지만, 본 연구에서는 Xplane을 입력받은 위치를 시각화하는 시각화 모듈로 사용하였다.

둘째, Virtual 시뮬레이터는 Constructive 시뮬레이터

에 전달해 주는 사용자의 반응을 입력해 주기 위한 조종석이 있다. VC 연동 시뮬레이터는 실험 환경을 현실감 있게 구성하는 것이 핵심이다. 현실감 있게 구성하지 않으면, 현실에서 얻을 수 있는 데이터 분포와 다른 분포로부터 실험 데이터를 추출하게 된다. Fig. 8은 이를 위한 조종석의 네 가지 하위 구성 요소를 나타내고 있다.

특히, Fig. 8(b)는 Flight Stick으로, 사용자의 Load Factor(N)와 Bank(P)를 입력받는 장치이다. Fig. 8(c)는 Throttle로, 사용자의 Thrust를 입력받는 장치이다. Fig. 8(d)는 시각화를 위한 스크린이다. 복수의 모니터를 연결하여 시각화할 경우, 현실과 괴리감 있는 실험 환경이 조성되기 때문에 본 연구에서는 부드러운 곡면을 가진, 현실과 유사한 환경을 조성하는 스크린을 사용하였다.

4. Human Experiment 실험 설계

위에서 논의한 바와 같이 본 논문의 시뮬레이션은 VC 연동 기반의 데이터 수집이 목적이다. 그러므로 모방 학습 및 역강화학습을 위한 데이터 수집 실험 설계는 중요하다. 이에 아래와 같이 사용자 실험을 설계하였다.

본 연구에서는 아군 기체(공격기)를 조종하여 적군 기체(방어기)를 격추시킨 데이터를 수집하였다. Table 2는 실험을 위한 기본 세팅을 나타낸다. 아군 기체의 초기 위치를 원점으로 하여, 아군 기체와 적군 기체의 거리 및 각도를 기준으로 적기 격추의 여부를 판단한다. 먼저, 첫 번째 격추 조건인, 아군 기체와 적군 기체의 거리 기준점은 800m로 설정하였다. 두 번째 격추 조건은 자세각으로, 지면 좌표계를 기준으로 한 피아 기체의 상대 벡터를 아군 기체 좌표계로 변환하여, 변환된 상대 벡터와 아군 기체 좌표계의 N방향의 각도로 정의한다. 격추 기준 각도는 실험 세팅에 따라 다르게 두었다.

기체 기동은 TNP 조종을 통해 이루어진다. 먼저, 회피

Table 2. Common experimental setting

	Initial blue location (North, East, Down) [m]	Shoot Distance [m]	T,N,P Control Range	Blue Speed Range [m/s]	Red Speed Range [m/s]	Max Acceleration [m/s ²]
Basic Setting	(0,0,0)	800	T 0-1 N 0.3-8.0 P -0.5-0.5	50-350	240-240	50

Table 3. Combat scenario by initial locations

	Initial Red Location (North/East/Down) [m]	Initial Red Attitude (Roll/Pitch/Yaw) [radian]	Number of Human Subjects	Number of Replication per Subjects	Shoot Bearing Anglet [radian]	Collection Data Trajectory
Scenario 1	(3000,0,0)	(0,0,0)	30	10	within 0.2	297
Scenario 2	(3000cos θ , 3000sin θ , 0) ¹⁾	(0,0, ϕ) ²⁾			within 0.2	257
Scenario 3	(3000,0,0)	(0,0,0)			within 0.1	79

를 하는 적기의 경우, 이전에 설명한 바와 같이 회피 기동을 위한 TNP 값을 설정한다. 또한, 아군 기체의 경우, 적기를 격추시키기 위하여 사용자가 입력한 TNP 값을 설정한다. 사용자는 Flight Stick 및 Throttle을 이용하여 TNP 신호를 Virtual 시뮬레이터에 전달하고, Virtual 시뮬레이터는 미리 지정된 TNP 범위로 전달된 신호를 정규화하여 Constructive 시뮬레이터에 전달한다. T의 범위는 0에서 1 사이로 정하고, N의 범위는 0.3에서 8 사이, P의 범위는 -0.5에서 0.5가 된다. T는 기체의 가속도를 결정하며, N과 P값은 기체의 선회 속도를 결정한다.

T가 0.5 이하이면 감속을, 0.5 이상이면 가속을 한다. N값이 1 이하이면 하강비행을, 1 이상이면 상승비행을 하며 자세각인 P값이 0 이하이면 기체를 왼쪽으로, 0 이상이면 기체를 오른쪽으로 회전시킨다.

아군 기체는 50m/s에서 350m/s 사이의 속력을 사용자의 입력을 받아 가지도록 세팅하였다. 적군 기체는 240m/s로 일정한 속력으로 회피 기동한다. 빠른 선회 기동을 위하여 가속도를 50m/s²으로 설정하였다.

Table 3은 실험 디자인 테이블이다. 본 연구에서는 총 세 가지 실험을 수행하였다. 총 30명의 실험자에 대하여 각각의 실험 세팅에 대하여 10개의 격추 데이터를 얻는 것을 목표로 하였다. 수집된 데이터는 시간에 따른 상태

벡터와 사용자 입력 TNP 벡터가 있다. 상태 벡터는 총 43개로, Table 4에 정리되어 있다.

첫 번째 실험은 고정된 위치에서 고정된 방향을 바라보게 초기화된 적기를 격추시키는 실험이다. 적기 격추 조건은 기체 간의 거리가 800m 이내이고 자세각이 0.2 radian 이내인 것으로 설정하여, 총 297개의 데이터를 수집하였다. 두 번째 실험은 적기가 일정 범위 내에서 움직이며 랜덤한 방향을 바라보게 초기화 된 세팅에서 적기를 격추시키는 실험이다. 피아 기체 간 거리가 800m 이내이고 자세각이 0.2 radian 이내일 때 적기가 격추되는 것으로 간주하여, 총 257개의 데이터를 수집하였다. 세 번째 실험은 첫 번째 실험과 동일하게 초기화한 세팅에서 격추 조건만 달리하여 진행하였다. 거리 조건은 동일하되, 자세각이 0.1 radian 이내일 때 적기가 격추되는 것으로 간주하여 실험을 진행하였고, 총 79개의 데이터를 수집하였다.

5. 데이터 수집 결과

Fig. 9는 시나리오 1에 대한 결과이다. Fig. 9(a)는 기체의 log-scale 거리와 그 평균을, (b)는 Red와 Blue Aircraft간의 bearing angle과 그 평균을 나타낸다. 격추 조건은 거리가 800m이내, 자세각이 0.2 radian 이내일 때로, 빨간색으로 표시된 평균이 시뮬레이션 시간이 진행됨에 따라 지속적으로 떨어지고 있는 것을 확인할 수 있다.

1) $\theta \sim U[-\frac{\pi}{3}, \frac{\pi}{3}]$

2) $\phi \sim U[-\pi, \pi]$

Table 4. State definitions recorded during experiments (which can be used in reinforcement learning)

No.	State Variables	Variable Dimensions
1	Player Position on NED coordinate (North, East, Down)	3
2	Player Position on Opponent coordinate (North, East, Down)	3
3	Opponent Position on NED coordinate (North, East, Down)	3
4	Opponent Position on Player coordinate (North, East, Down)	3
5	Player Attitude on NED coordinate (Roll, Pitch, Yaw)	3
6	Player Attitude on Opponent coordinate (Roll, Pitch, Yaw)	3
7	Opponent Attitude on NED coordinate (Roll, Pitch, Yaw)	3
8	Opponent Attitude on Player coordinate (Roll, Pitch, Yaw)	3
9	Player Velocity on NED coordinate (Velocity of North, East, Down)	3
10	Player Velocity on Opponent coordinate (Velocity of North, East, Down)	3
11	Opponent Velocity on NED coordinate (Velocity of North, East, Down)	3
12	Opponent Velocity on Player coordinate (Velocity of North, East, Down)	3
13	Player-Opponent Distance	1
14	Line of Sight Angle from Player to Opponent (Yaw, Pitch)	2
15	Line of Sight Angle from Opponent to Player (Yaw, Pitch)	2
16	Bearing Angle from Player to Opponent	1
17	Bearing Angle from Opponent to Player	1
	Total	43

특히, 시뮬레이션이 진행됨에 따라, 격추 조건에 점진적으로 평균이 접근해 가는 것을 확인할 수 있다. Fig. 9(c)는 아군 좌표계에서 바라본 적군의 Pitch와 평균을, Fig. 9(d)는 아군 좌표계에서 바라본 적군의 Yaw와 평균을 나타낸다. Fig. 9(c)에서 실험자의 시야는 오직 아군 기체 이동 방향으로부터 위 방향만 관측 가능하기 때문에, 실험자는 지속적으로 적군 기체를 올려다보며 추격한다. 이는 자연 추적 방법을 의미하는데, 시간이 지남에 따라 빨간색 평균 Pitch 값이 내려가는 것은 실험자가 자연 추적

에서 벗어나 적기를 잡기 위해 일치 추적 및 선도 추적으로 전환하였음을 나타낸다. Fig. 9(d)에서 아군 좌표계에서 바라본 적군의 Yaw가 $\pm \pi$ 근처 값을 갖는 것은 아군이 적군을 시야에서 놓친 것을 의미하며, 적군의 회피 기동이 성공적으로 작동하고 있다는 것을 의미한다. 빨간색 평균과 보라색 선은 각각 적군의 초기 회피 경로가 오른쪽/왼쪽으로 이동한 에피소드에 대한 Yaw의 평균이다. 각각 5초에서 15초 사이로 경로를 왼쪽에서 오른쪽으로, 혹은 오른쪽에서 왼쪽으로 틀며 이동을 하는 것을 관측할 수 있다.

Fig. 10은 시나리오 2에 대한 결과를 나타낸다. Fig. 10은 적기의 초기 위치 및 방향이 변하기 때문에 자세각을 나타낸 (b)에서 자세각이 0으로 초기화되어있지 않다. 적기의 초기 위치는 $[-\frac{\pi}{3}, \frac{\pi}{3}]$ 에서 움직이기 때문에 초기 자세각은 $[0, \frac{\pi}{3}]$ 안에 위치하고, Fig. 10(b)에서 우리는 초기 자세각이 $\frac{\pi}{6} \approx 0.523598$ 근처에 있는 것을 확인할 수 있다. 위치는 변하지만 초기 거리는 변하지 않기 때문에 (a)에서 거리는 3000m로 일정하다. 마찬가지로, Yaw 또한 초기에 0으로 고정되어 있지 않기 때문에 Fig. 10(d)에서 초기 평균값이 0이 아닌 것을 확인할 수 있다.

Fig. 11은 시나리오 3의 결과를 나타낸다. 시나리오 3은 격추하기 위한 자세각 조건(0.1 radian 이내)을 만족하기 어렵기 때문에 상대적으로 시나리오 1과 2에 비해 적은 데이터를 수집하였다. Fig. 11은 시나리오 1과 마찬가지로 초기화되어있고, 격추 조건만 다르기 때문에 유사한 통계 결과를 나타낸다.

Fig. 12는 각 실험 세팅에 대한 적군 기체를 격추한 격추 시간에 대한 분포이다. 시나리오 3이 가장 어려운 실험으로, Fig. 12(c)는 시나리오 3의 heavy tailedness를 보여준다. 시나리오 1은 비교적 시나리오 2보다 더 높은 비율로 빠른 시간 안에 종료되지만, 한번 적기를 격추할 기회를 놓칠 경우, 시나리오 2보다 더 tail에 많은 밀도가 결집되어 있는 것을 확인할 수 있다.

3개의 시나리오에 대한 실험을 통해 사용자(blue)와 적기(red)의 상황 및 그에 대한 행동에 대한 상태 벡터 (Table 4 참조)와 시뮬레이션의 결과로 도출된 kill time 정보를 확인할 수 있다(Fig. 12 참조). 획득한 데이터는 역강화학습의 관점에서 볼 때 학습에 필요한 에이전트의 행동(action), 상황(state) 그리고 그에 대한 보상(reward)에 대한 데이터로써 활용이 가능하다.

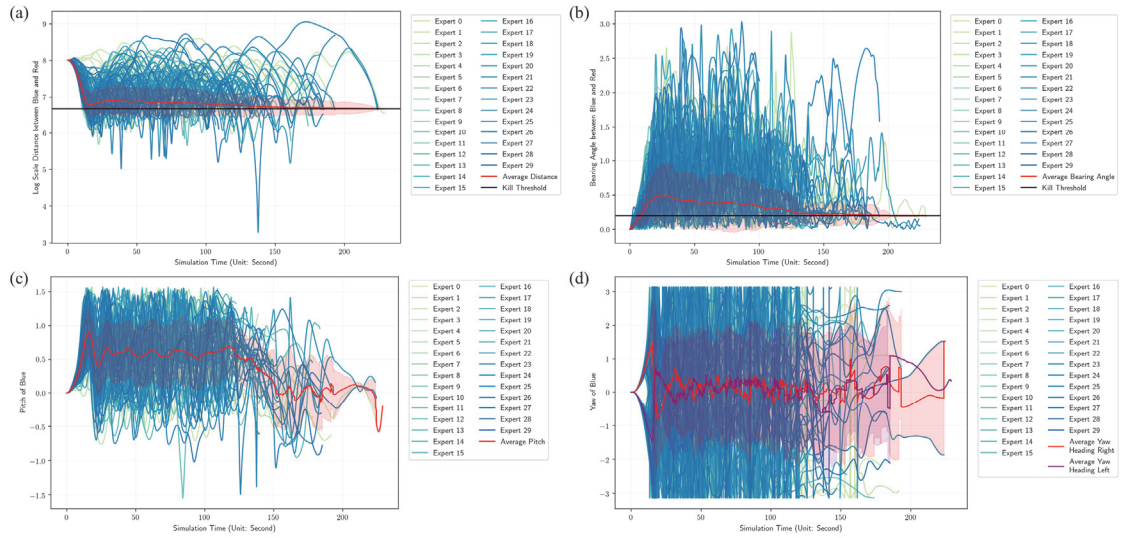


Fig. 9. Scenario 1 (a) Log-scale distance between two aircrafts over time (b) Bearing angle between two aircrafts over time (c) Pitch angle of blue aircraft over time (d) Yaw angle of blue aircraft over time

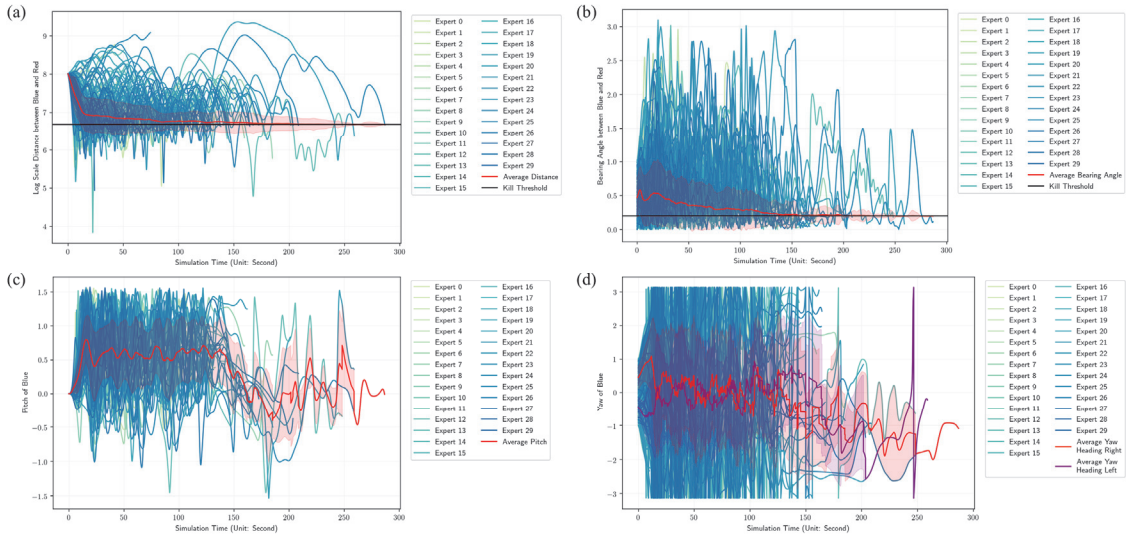


Fig. 10. Scenario 2 (a) Log-scale distance between two aircrafts over time (b) Bearing angle between two aircrafts over time (c) Pitch angle of blue aircraft over time (d) Yaw angle of blue aircraft over time

6. 결론

본 논문은 역강화학습을 위한 사용자 데이터 수집을 VC연동을 통해 수행하는 사례를 알려준다. 역강화학습은 사람과 같이 행동하는 인공지능을 만들기 위한 기초적인 방법이다. 최근의 지도기반 기계학습의 발전으로 다

양한 이미지 혹은 텍스트 기반의 데이터가 사용자에게 의해 구성되고 있다. 본 논문은 제안된 VC연동 시뮬레이션이 역강화학습 혹은 모방학습을 위한 학습 데이터의 구성에 활용될 수 있음을 시뮬레이션 결과로 획득한 데이터를 통해 시사하고 있다. 이러한 과정을 통해, VC 연동 시뮬레이션이 인공지능 분야에 더욱 활발하게 활용될 수

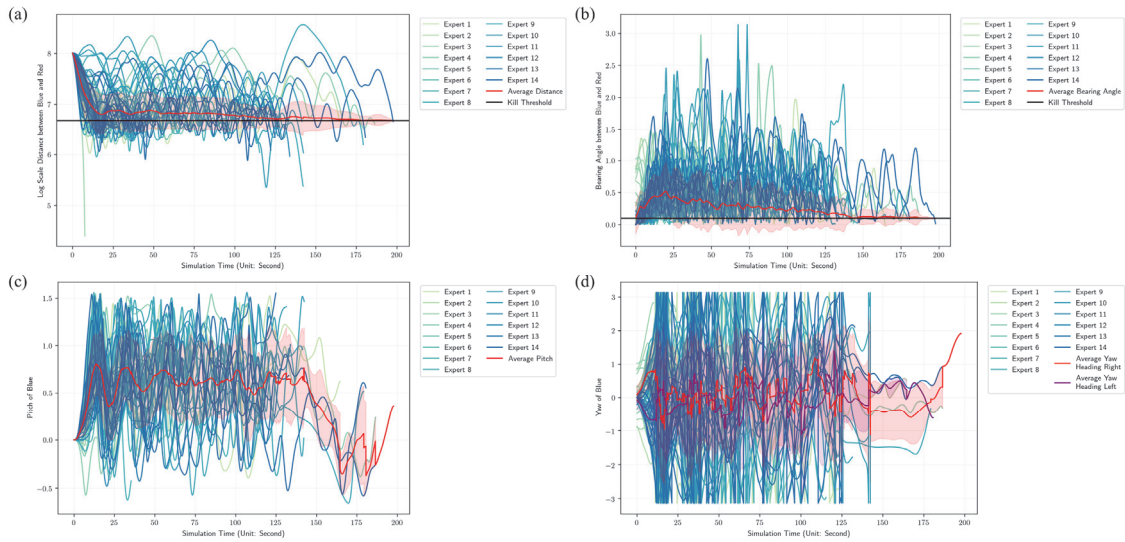


Fig. 11. Scenario 3 (a) Log-scale distance between two aircrafts over time (b) Bearing angle between two aircrafts over time (c) Pitch angle of blue aircraft over time (d) Yaw angle of blue aircraft over time

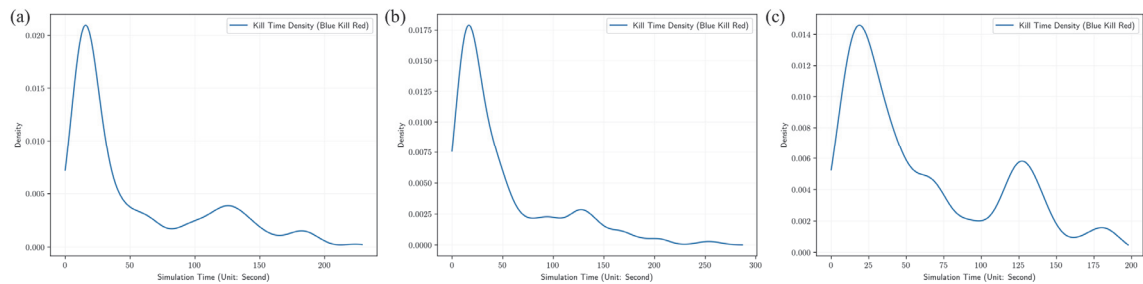


Fig. 12. Simulation elapse time until the blue shoot down the red aircraft (a) Scenario 1, (b) Scenario 2, (c) Scenario 3

있기를 기대한다.

References

- 이동진, 방효충 (2009), “강화학습을 이용한 무인전투기 (UCAV) 근접 공중전”. *한국항공우주학회 학술발표회 초록집*, pp. 249-252.
- 문일철, 정민재, 김동준 (2020), “깊은강화학습 기반 1-vs-1 공중전 모델링 및 시뮬레이션”. *한국시뮬레이션학회 논문지*, 29(1), pp. 39-46.
- 한승엽, 이병운, 탁민재 (2016), “Scoring Function Matrix를 활용한 2:1 전투기 3차원 공중전 기동 생성”. *한국항공우주학회 학술발표회 초록집*, pp. 435-436.
- Abbeel, P., & Ng, A. Y (2004), “Apprenticeship learning

via inverse reinforcement learning”, *In Proceedings of the twenty-first international conference on Machine learning*, pp. 1.

- Bonanni, Pete (1993), *The art of the kill*. Spectrum HoloByte
- Ernest, N., Carroll, D., Schumacher, C., Clark, M., Cohen, K., & Lee, G (2016), “Genetic fuzzy based artificial intelligence for unmanned combat aerial vehicle control in simulated air combat missions”. *Journal of Defense Management*, 6(144), pp. 2167-0374.
- Ho, Jonathan, and Stefano Ermon (2016), “Generative adversarial imitation learning.”, *Advances in neural information processing systems(NIPS)*. 2016.

Toghiani-Rizi, B., Kamrani, F., Luotsinen, L. J., & Gisslén, L (2017), "Evaluating deep reinforcement learning for computer generated forces in ground combat simulation". In *Proceedings of 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 3433-3438

Sutton, R. S., & Barto, A. G (1998), *Introduction to reinforcement learning*. Cambridge: MIT press.
Shaw, R. L (1985), *Fighter Combat*. Naval Institute Press.
Zeigler, Bernard P., Tag Gon Kim, and Herbert Praehofer, (2000), *Theory of Modeling and Simulation*.



김 동 준 (ORCID : <https://orcid.org/0000-0002-1117-5848> / dongjoun57@kaist.ac.kr)

2016 한국과학기술원 수리과학과 학사
2016~ 한국과학기술원 산업및시스템공학과 석박사 통합과정

관심분야 : 인공지능, 시뮬레이션, 모델링



신 용 진 (ORCID : <https://orcid.org/0000-0002-5961-933X> / yj.shin@kaist.ac.kr)

2019 포항공과대학교 산업경영공학과 학사
2019~ 한국과학기술원 산업및시스템공학과 석사

관심분야 : 인공지능, 분산학습



안 경 수 (ORCID : <https://orcid.org/0000-0002-7151-8339> / kyeongsu.an@hanwha.com)

1996 부산대학교 전자전기통신공학부 학사
2009 경북대학교 전기전자컴퓨터공학부 석사
2002~ 한화시스템 항공우주연구소 항공시스템팀 팀장



김 영 곤 (ORCID : <https://orcid.org/0000-0001-6348-9980> / younggon.kim@hanwha.com)

2002 경일대학교 컴퓨터공학과 학사
2009 포항공과대학교 정보통신학과 석사
2002~ 2016 (주)휴원 국방항공사업부 부장
2016~ (주)한화시스템 항공우주연구소 항공시스템팀 수석연구원



문 일 철 (ORCID : <https://orcid.org/0000-0002-1798-1306> / icmoon@kaist.ac.kr)

2004 서울대학교 컴퓨터공학부 학사
2005 카네기멜론대학교, Information Systems Management, Heinz Public Policy School, 석사
2008 카네기멜론대학교 전산학 박사
2008~ 2011 한국과학기술원 SMSLab 박사후연구원
2011~ 한국과학기술원 산업 및 시스템공학과 교수

관심분야 : 모델링, 시뮬레이션, 인공지능



배 장 원 (ORCID : <https://orcid.org/0000-0002-4681-5059> / jangwon_bae@koreatech.ac.kr)

2007 고려대학교 전기전자전파공학부 학사
2009 한국과학기술원 전자공학과 석사
2015 한국과학기술원 산업및시스템공학과 박사
2015~ 2020 한국전자통신연구원 인공지능연구소 선임연구원
2020~ 한국기술교육대학교 산업경영학부 조교수

관심분야 : 모델링, 시뮬레이션, 인공지능