

논문 2024-19-17

심층 강화학습 기반 자율운항 CTV의 해상풍력발전단지 내 장애물 회피 시스템

(Obstacle Avoidance System for Autonomous CTVs in Offshore Wind Farms Based on Deep Reinforcement Learning)

김진균, 전해명, 노재규*
(Jingyun Kim, Haemyung Chon, Jackyou Noh)

Abstract : Crew Transfer Vessels (CTVs) are primarily used for the maintenance of offshore wind farms. Despite being manually operated by professional captains and crew, collisions with other ships and marine structures still occur. To prevent this, the introduction of autonomous navigation systems to CTVs is necessary. In this study, research on the obstacle avoidance system of the autonomous navigation system for CTVs was conducted. In particular, research on obstacle avoidance simulation for CTVs using deep reinforcement learning was carried out, taking into account the currents and wind loads in offshore wind farms. For this purpose, 3 degrees of freedom ship maneuvering modeling for CTVs considering the currents and wind loads in offshore wind farms was performed, and a simulation environment for offshore wind farms was implemented to train and test the deep reinforcement learning agent. Specifically, this study conducted research on obstacle avoidance maneuvers using MATD3 within deep reinforcement learning, and as a result, it was confirmed that the model, which underwent training over 10,000 episodes, could successfully avoid both static and moving obstacles. This confirms the conclusion that the application of the methods proposed in this study can successfully facilitate obstacle avoidance for autonomous navigation CTVs within offshore wind farms.

Keywords : Deep Reinforcement Learning, Crew Transfer Vessel, Offshore Wind Farm, Convention on the International Regulations for Preventing Collisions at Sea, Obstacle Avoidance

1. 서론

현재 인류가 직면하고 있는 문제 중 가장 심각한 위기는 지구 온난화로 인한 기후변화이다. 세계 각 나라와 기업들은 이를 막기 위해 탄소 배출량을 줄이고자, 2050 탄소중립(NET-ZERO)과 RE100 (Renewable Electricity 100%) 캠페인에 참여하고 있다. 특히, 기업들의 RE100 목표 달성과 관련하여 재생에너지로 생산된 전기를 사용하기 위해서는 태양광과 풍력 에너지 발전의 발전 용량이 현 발전 용량 대비 크게 증설되어야 하며, 이를 위해서는 대규모 해상풍력발전단지의 신규 설치가 불가피한 상황이다.

대규모 해상풍력발전단지가 설치된 이후에는 이에 대한 유지관리와 보수가 필요하며, 우리나라는 해안선이 복잡하여 해상풍력발전기의 고장률이 높다 [1]. 해상풍력발전기의 유지관리 및 보수에는 주로 유지보수 작업인력 수송선인 CTV (Crew Transfer Vessel)가 사용된다. CTV는 해상풍력발전단지 내에서 항해할 때, 기상, 조류, 파도와 같은 해상 환경의 영향을 받고, 해상풍력발전기까지의 접근성이 용이

하지 않기 때문에 전문 선장과 승무원이 직접 조종한다. 그럼에도 불구하고 CTV는 높은 운항 난이도로 인해 운항 중 타 선박이나 해양구조물과 같은 장애물과 충돌하는 사고가 발생한다.

선박 충돌사고의 원인에는 경계 소홀과 항해 법규 위반과 같은 인적 요인이 큰 비중을 차지하고 있다. 인적 요인에 의해 발생하는 충돌사고를 방지하기 위해서는 선박에 자율운항시스템을 도입할 필요가 있다 [2].

선박에 자율운항시스템을 도입하기 위한 기술 중 장애물 회피 연구를 위해서는 선박의 동역학적 특성을 반영한 선박 조종 운동 모델링이 수행되어야 하며, 장애물 충돌 여부 및 위험도 판단 기법과 장애물 회피 기동 제어 알고리즘에 관한 연구가 수행되어야 한다. 또한, 해상풍력발전단지 내 CTV 운영을 위해 해상풍력발전단지과 유사한 외부 환경 하중을 고려한 시뮬레이션 환경 구현도 필요하다.

많은 선박의 장애물 회피 관련 연구에서는 선박 조종 운동 모델링 수행 시, 선미 좌우에 2개의 추진기 (Thruster) 간 출력 차이를 이용해 조종하는 무인수상선 (USV, Unmanned Surface Vessel) [3], 선미 좌우에 2개의 프로펠러 간 출력 차이를 이용해 조종하는 무인수상선 [4-6] 그리고 2개의 추진기와 조타기를 이용해 조종하는 무인수상선을 대상 선박 모델로 사용하였다 [7].

*Corresponding Author (snucurl@kunsan.ac.kr)
Received: Feb. 20, 2024, Revised: Apr. 4, 2024, Accepted: Apr. 8, 2024.
J. Kim: Kunsan National University (Ph.D. candidate)
H. Chon: Kunsan National University (Ph.D. candidate)
J. Noh: Kunsan National University (Prof.)

장애물 충돌 여부 판단 기법으로는 많은 연구에서 국제해상충돌예방규칙 (COLREGs, Convention on the International Regulations for Preventing Collisions at Sea) 를 고려하였으며, 선박의 안전 영역 기반 충돌 위험도 판단 기법 [3, 7], CPA (The Closest Point of Approach) 기반 충돌 여부 판단 기법 [7], 심층 강화학습 기반 충돌 여부 판단 기법 [5, 6]을 사용한 연구가 수행되었다. 장애물 회피 알고리즘은 Fuzzy 제어를 사용한 변침각 제어 알고리즘 [4], 심층 강화학습을 사용한 장애물 회피 경로 생성 알고리즘 [5, 6], 심층 강화학습을 사용한 조타각 제어 알고리즘 [7]과 관련한 연구가 수행되었다.

이처럼 선박의 장애물 회피에 관한 연구는 꾸준히 수행되었다. 하지만, 실제 해상풍력발전단지 유지보수에 사용되는 워터젯 추진 선박인 CTV의 조종 운동 모델을 사용한 선박의 장애물 회피 연구와 해상풍력발전단지와 유사한 외부 환경을 고려한 시뮬레이션을 구현한 선박의 장애물 회피 연구는 부족한 상황이다.

따라서, 본 연구에서는 해상풍력발전단지와 유사한 환경에서 CTV의 장애물 회피에 관해 연구하기 위해 무인수상선의 대상 선박을 워터젯 추진 선박으로 선정하였다. 선박 조종 운동 모델링 중 조류 하중은 향후 해상풍력발전단지가 설치될 수 있는 군산 앞바다의 2021년 연간 조류 데이터를 기반으로 하중을 구현하였으며, 풍하중은 이격거리에 따른 해상풍력발전단지 최적 설계에 관한 연구 [8] 중 서해안의 영광군 앞바다에 위치한 왕등여 지역에서 2011년 1월부터 2011년 12월까지 관측한 실측풍속 데이터를 기반으로 풍하중을 구현하였다. 장애물 회피 시뮬레이션 시험 환경은 해상풍력발전단지와 유사한 환경을 구현하기 위해 위와 같은 연구 [8]를 참고하여 2.5km * 4km 크기의 단지를 가정하고 시뮬레이션 시험 환경을 구현하였다. 또한, 심층 강화학습은 고차원의 입력 데이터와 고차원의 출력 데이터를 다룰 때 성능이 좋은 MATD3 (Multi-Agent Twin Delayed Deep Deterministic Policy Gradient)를 사용해 장애물 회피 기동에 관한 연구를 수행하였다.

본 연구의 2장에서는 워터젯 추진 선박의 조종 운동 모델링에 관한 내용을 기술하였으며, 3장에서는 심층 강화학습 에이전트의 학습 및 시험에 관한 내용을 기술하였다.

II. 선박 조종 운동 모델

1. 선박 조종 운동 모델

선박은 기본적으로 6 자유도 운동을 한다. 선박의 조종 운동을 표현하기 위해 6 자유도 운동을 고려하여 모델링을 하게 되면 실제 선박과 매우 유사한 움직임을 표현할 수 있다. 하지만 6 자유도 운동을 기반으로 모델링 하게 되면 이를 표현하기 위한 수식이 복잡해지고, 복잡한 시스템이 요구되기 때문에 실제 모델링에 어려움이 있다 [9]. 따라서, 본 연구에서는 선박을 6 자유도 운동 중 조종 성능과 밀접한 관련이 있는 전후 동요 (Surge), 좌우 동요 (Sway), 선수동요

표 1. 쌍동선형 선박의 주요 제원 [5]

Table 1. Principle specification of a Catamaran Vessel [5]

Parameter	Value
Length Over All (LOA)	4.88m
Beam	2.5m
Draft	0.3m
Mass	151kg
Moment of Inertia (I_{ZZ})	257kg·m
Length from Perpendicular to Perpendicular (LBP)	4.58m
Frontal projected area (A_{FW})	1.53m ²
Lateral projected area (A_{LW})	2.31m ²

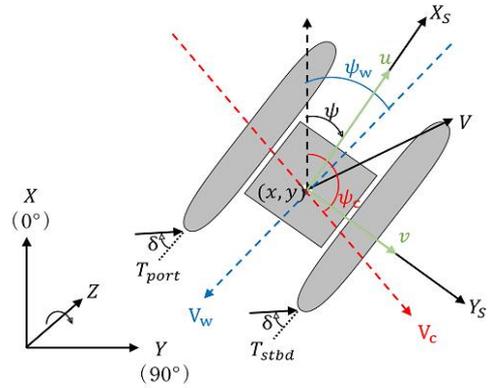


그림 1. 선박 조종 운동 모델의 기준 좌표계

Fig. 1. Coordinate system for ship maneuvering model

(Yaw) 총 3개의 운동 성분을 가진 3 자유도 수평면 운동을 하는 강체 (Rigid body)라 가정하고 연구를 수행하였다.

1.1 무인 워터젯 추진 선박 운동 모델

본 연구에서는 심층 강화학습 기반 장애물 회피 관련 연구 [5]를 참고하여 표 1과 같은 제원을 가진 쌍동선형 선박의 운동 모델을 기반으로 그림 1과 같은 좌표계를 사용하여 워터젯 추진 기반 무인선의 조종 운동을 표현하였다. 그림 1은 연구에서 사용된 좌표계를 나타낸 그림이다. 그림에서 X 는 지구 고정 좌표계의 북쪽 및 0° , Y 는 지구 고정 좌표계의 동쪽 및 90° 를 의미하며 x 와 y 는 지구 고정 좌표계 상에서 선박의 위치를 의미한다. V_c 와 V_w 는 지구 고정 좌표계 상에서 조류와 바람의 속도, ψ , ψ_c 그리고 ψ_w 는 지구 고정 좌표계 상 Y_s 에서 선수, 조류 그리고 바람의 방향을 의미한다. X_s 와는 선체 중심 고정 좌표계, v , u 그리고 v 는 선체 중심 고정 좌표계에서 전속도, 선수 방향의 속도, 우현 방향의 속도를 의미한다. δ , T_{port} , T_{stbd} 는 선체 중심 고정 좌표계에서 각 워터젯의 출력 방향, 좌현과 우현 워터젯의 출력을 의미한다.

선박의 동역학적 특성에 관한 연구 [10]에 따르면 지구 고정 좌표계에서 선박의 위치 및 자세 벡터를 $\eta = [x, y, \psi]$, 선체 중심 고정 좌표계에서 선박의 전속도 및 각속도 벡터를 $\nu = [u, v, r]$ 라고 할 때, 선박의 동역학

방정식은 식 1과 같이 정의할 수 있다. 식 1에서 $R(\psi)$ 은 주성분 회전 매트릭스 (Principle rotation matrix)를 의미한다. 3 자유도 운동을 하는 선박의 비선형 동역학 운동 방정식은 식 2와 같이 표현할 수 있다.

식 2에서 M 은 부가 질량 관성 행렬, $C(\nu)$ 는 부가 질량 코리올리 (Coriolis)와 구심 항 (Centripetal terms) 행렬, $D(\nu)$ 는 감쇠 행렬 (Damping matrix), f 는 제어 입력 벡터를 의미한다. 본 연구에서 제어 입력 벡터 f 는 식 3과 같으며, 여기서 τ_x, τ_y 는 각 축 방향 힘, τ_N 는 중심점 모멘트를 의미한다.

식 2와 식 3을 정리하고, 동역학이 좌우 동요 및 선수동요와 비연성 되었다고 가정하고 선형화를 진행하면, 해당 선박의 조종 운동 방정식을 의미하는 식 4를 구할 수 있다. 식 4는 무인수상선의 동역학을 표현하는 선형시불변 (Linear time invariant) 상태 공간 (State space) 방정식으로 $a_{11}, a_{22}, \dots, a_{33}$ 는 선박의 동역학 특성을 나타내기 위한 미지수, b_{11}, b_{22} 그리고 b_{32} 는 선박에 조종 명령에 따라 발생하는 힘과 모멘트의 특성을 나타내기 위한 미지수이다.

$$\dot{\eta} = R(\psi)\nu, \quad R(\psi) = R_{z,\psi} = \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (1)$$

$$M\dot{\nu} + C(\nu)\nu + D(\nu)\nu = f, \quad (2)$$

$$f = \begin{bmatrix} \tau_x \\ \tau_y \\ \tau_N \end{bmatrix} = \begin{bmatrix} T_{port}\cos(\delta) + T_{stbd}\cos(\delta) \\ T_{port}\sin(\delta) + T_{stbd}\sin(\delta) \\ (-T_{port}\sin(\delta) + T_{stbd}\sin(\delta))B/2 \end{bmatrix}, \quad (3)$$

$$\begin{bmatrix} \dot{u} \\ \dot{v} \\ \dot{r} \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} u \\ v \\ r \end{bmatrix} + \begin{bmatrix} b_{11} & 0 \\ 0 & b_{22} \\ 0 & b_{32} \end{bmatrix} \begin{bmatrix} \tau_x \\ \tau_N \end{bmatrix}. \quad (4)$$

1.2 풍하중 운동 모델

선체에 작용하는 풍하중 운동 모델은 식 5와 같이 풍하중 내 선박의 동역학적 특성에 관한 연구 [10]의 풍하중 운동 모델을 사용하였다.

$$\tau_w = \begin{bmatrix} \tau_{wx} \\ \tau_{wy} \\ \tau_{wN} \end{bmatrix} = \frac{1}{2}\rho_a V_{rel}^2 \times \begin{bmatrix} C_X(\psi_{rel})A_{FW} \\ C_Y(\psi_{rel})A_{LW} \\ C_N(\psi_{rel})A_{LW}LOA \end{bmatrix}. \quad (5)$$

식 5에서 ρ_a 는 공기밀도 (1.2 kg/m^3), V_{rel} 는 선체에 대한 바람의 상대속력, ψ_{rel} 는 선체에 대한 바람의 상대 방향, $C_X(\psi_{rel})$, $C_Y(\psi_{rel})$ 그리고 $C_N(\psi_{rel})$ 은 풍하중 계수, A_{FW} 는 전 투영 면적 (Frontal projected area), A_{LW} 는 측 투영 면적 (Lateral projected area) 그리고 LOA 는 전장을 의미한다. V_{rel} 와 ψ_{rel} 는 식 6 ~ 식 9을 이용해 구할 수 있다.

$$u_{rel} = V_w \cos(\psi_w - \psi) + u. \quad (6)$$

$$v_{rel} = V_w \sin(\psi_w - \psi) + v. \quad (7)$$

$$V_{rel} = \sqrt{u_{rel}^2 + v_{rel}^2}, \quad (8)$$

$$\psi_{rel} = \tan^{-1}\left(\frac{v_{rel}}{u_{rel}}\right). \quad (9)$$

풍하중 계수 $C_X(\psi_{rel})$, $C_Y(\psi_{rel})$ 그리고 $C_N(\psi_{rel})$ 는 식 10 ~ 식 12를 이용해 근사값을 구할 수 있다.

$$C_X(\psi_{rel}) = -c_X \cos(\psi_{rel}), \quad (10)$$

$$C_Y(\psi_{rel}) = c_Y \sin(\psi_{rel}), \quad (11)$$

$$C_N(\psi_{rel}) = c_Z \sin(2\psi_{rel}). \quad (12)$$

식 10 ~ 식 12에서 c_X , c_Y , c_Z 는 식 13에 속하는 값으로 근사화할 수 있다. 본 연구에서는 풍하중 모형실험을 수행할 때, 같은 선박 모델을 사용하여 풍하중 계수를 근사화한 연구 [11]를 참고하여 식 14와 같이 정의하였다.

$$c_X \in (0.5, 0.9), \quad c_Y \in (0.7, 0.95), \quad c_Z \in (0.05, 0.2), \quad (13)$$

$$c_X = 0.7, \quad c_Y = 0.8, \quad c_Z = 0.1. \quad (14)$$

1.3 조류 하중 운동 모델

본 연구에서 조류 하중 운동 모델은 선체에 작용하는 조류 하중에 관한 연구 [12]를 참고하여 식 15 ~ 식 17로 표현되는 운동 모델을 사용하였다.

$$X_c = m(\dot{u}_r - v_r r - x_y r^2), \quad (15)$$

$$Y_c = m(\dot{v}_r + u_r r - x_c r^2), \quad (16)$$

$$N_c = I_{ZZ}(\dot{r} + m x_c (\dot{v}_r + u_r r)). \quad (17)$$

식 15 ~ 식 17에서 X_c , Y_c 그리고 N_c 는 각 지구 고정 좌표계 상에서 X 방향, Y 방향 그리고 Z 방향으로 조류가 미치는 영향을 의미한다. u_r 과 v_r 는 선체에 작용하는 조류의 상대속력, \dot{u}_r 과 \dot{v}_r 는 가속도, r 은 선수의 각속도, \dot{r} 은 선수의 각 가속도, m 은 선체의 질량, I_{ZZ} 는 질량 관성 모멘트를 의미한다.

선체에 작용하는 조류의 상대속력 u_r 과 v_r 는 식 18과 식 19와 같이 구할 수 있다.

$$u_r = u - V_c \cos(\psi - \psi_c), \quad (18)$$

$$v_r = v + V_c \sin(\psi - \psi_c). \quad (19)$$

2. 선박 조종 운동 시뮬레이션

선박 조종 운동 모델이 제대로 설계되었는지 확인하기 위해 실제 운항하고 있는 선박의 운동과의 결과 비교를 통한 성능 평가가 필요하지만, 실제 선박이 통제된 바람, 조류 환경에서 시험하는 것은 현실적인 어려움이 있어 운동 모델링에 대한 시뮬레이션 수행을 통하여 운동 모델의 타당성을 검증하였다.

선박 조종 운동 시뮬레이션은 선회 시뮬레이션을 진행하였으며, 선박의 속도를 10m/s로 고정하고 다음과 같은 실험 환경을 가정하여 진행하였다.

- 환경 하중은 존재하지 않으며, 워터젯의 출력 방향을 10°, 20°, 45°로 유지 (그림 2 (a) ~ (c))

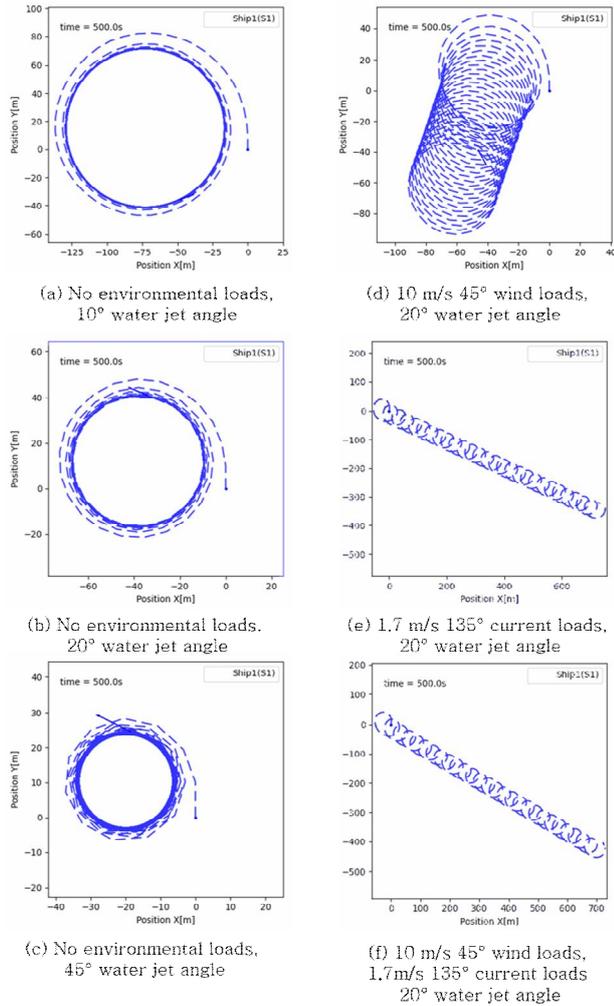


그림 2. 선회 시뮬레이션 결과
Fig. 2. Result of turning circle test

- 10m/s, 45° 방향의 풍하중만 존재, 워터젯의 출력 방향을 20°로 유지 (그림 2 (d))
- 1.7 m/s, 135° 방향의 조류 하중만 존재, 워터젯의 출력 방향을 20°로 유지 (그림 2 (e))
- 10m/s, 45° 방향의 풍하중과 1.7m/s, 135° 방향의 조류 하중에 동시에 존재, 워터젯의 출력 방향을 20°로 유지 (그림 2 (f))

그림 2는 선회 시뮬레이션 결과이며, 각 그래프는 다양한 환경 조건 하에서의 선박 선회 경로를 의미한다. 모두 원점 (0, 0)에서 선수가 북쪽 (0°)을 바라본 상황에서 시뮬레이션을 시작하였다.

선회 시뮬레이션 결과 그림 2 (a) ~ (c)와 같이 환경 하중이 없을 때, 선박은 일정한 반경의 원형 선회 경로를 그리며 안정적인 선회 수렴성을 가지는 것을 확인할 수 있다. 하지만, 그림 2 (d) ~ (f)와 같이 외부 환경 하중이 존재하면, 선박의 선회 경로는 강한 풍하중과 조류 하중에 의해 크게 영향을 받으며, 특히, 조류 하중의 영향을 크게 받는다는 것을 확인할 수 있다.

III. 심층 강화학습 기반 장애물 회피 시스템

1. 장애물 회피 의사결정

선박이 목적지까지 운항할 때, 장애물을 회피하는 절차는 전역 경로 생성 - 전역 경로 추종 - 장애물 탐지 - 장애물 회피 의사결정 - 장애물 회피 기동으로 구성된다. 본 연구에서는 장애물 회피 절차 중 장애물 회피 기동할 때 심층 강화학습을 적용하여 장애물을 회피 및 기동하는 알고리즘에 관해 연구하였다. 따라서, 장애물 회피 의사결정 여부는 선박이 장애물을 탐지했을 때 장애물과 충돌 위험이 있다면 반드시 장애물을 회피하도록 시스템을 구성하였다. 시스템은 고정 장애물과 이동 장애물을 구분하여 충돌 위험을 판단하고, 장애물 회피 여부를 결정한다.

선박의 경로는 임의의 출발지 A 지점과 임의의 목적지 B 지점까지 연결한 직선을 따라 이동한다고 가정한다. 또한, 해상에서 선박이 장애물의 위치와 자세를 추정하기 위해서는 AIS (Aeronautical Information Services), Radar, LiDAR 등의 센서와 장비를 이용하지만, 본 연구에서는 선박이 장애물에 대한 위치와 자세를 알고 있다고 가정하였다.

1.1 고정 장애물의 충돌 위험 판단

고정 장애물은 해상에서 바람과 조류의 영향을 받지 않아 고정된 위치에 있을 수 있다고 가정하였다. 본선이 고정 장애물을 회피하기 위해서는 장애물의 위치와 크기가 필요하다. 본 연구에서의 고정 장애물은 모두 20m 와 40m 직경의 원통형 구조물로 가정하였으며, 장애물 탐지 시 다음 조건과 같이 충돌 위험을 판단하였다.

- 본선과 장애물의 사이 각과 본선이 바라보고 있는 각의 차이가 13° 이하를 만족하면서 사이 거리가 300m 이하일 경우: 충돌 회피 기동 필요
- 본선과 장애물 사이 각과 본선이 바라보고 있는 각의 차이가 45° 이하를 만족하면서 사이 거리가 장애물과 거리가 200m 이하일 경우: 충돌 회피 기동 필요
- 본선과 장애물 사이 거리가 300m 이하를 만족하면서, 위 두 항에 언급된 사이 각 조건을 만족하지 않거나, 본선과 장애물 사이 거리가 300m 이상을 만족하는 경우: 기존 항로 유지

1.2 COLREGs 규정과 이동 장애물의 충돌 위험 판단

COLREGs는 항해 중 각 조우 상황에 대하여 선박을 피항의무선 (Give-way vessel) 과 진로유지선 (Stand-on vessel)으로 구분하여 조치해야 할 절차에 관해 다루고 있다. 본 연구에서는 300m 이내에 들어오는 이동 장애물에 대해 주요 COLREGs 규칙을 기반으로 장애물 회피 의사결정에 관여하여 다음과 같이 규정하였다.

다음은 본 연구에서 참고한 COLREGs 규칙을 나타내며, 그림 3은 각 COLREGs 규칙을 도식화한 그림이다. 그림 3에서 검은색 삼각형은 본선 (Own Ship, OS), 빨간색 삼각형은 타선 (Target Ship, TS), 파란색 영역은 본선의 선수 각 영역, 노란색은 타선의 선수 각 영역을 의미한다.

- 제14조. 정면으로 마주치는 상태 (Head-on situation)

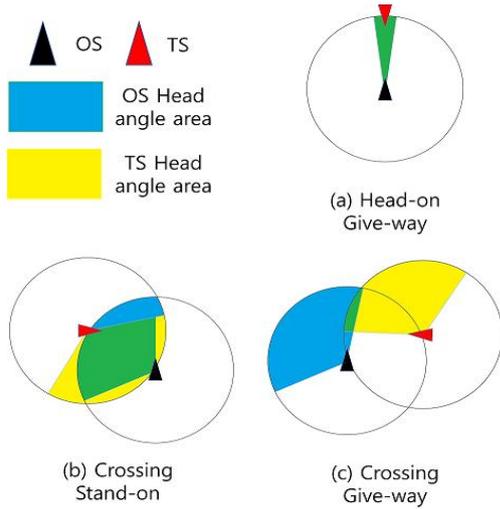


그림 3. 피항의무선과 진로유지선의 구분
Fig. 3. Classification of give-way vessel and stand-on vessel

두 척의 동력선이 서로 마주친 경우 두 척의 동력선 모두 타 선박의 좌현 측을 통과할 수 있도록 우현 측으로 항로를 변경하여야 한다.

: 본선의 선수각과 타선 간 사잇각. 그리고 본선 선수각과의 차이가 6° 미만인 경우 :: 피항 의무선 (그림 3 (a))

- 제15조. 교차 상태 (Crossing situation)

교차 상태에서는 피항의무선과 진로유지선의 관계가 성립된다. 그리고 자신의 우현 측에 타선을 보는 선박은 피항의무선이 조치를 취해야 하며, 상황이 허락하는 한 진로유지선의 선수를 지나가는 행동은 피해야 한다.

: 본선의 선수각과 타선 간 사잇각, 그리고 본선 선수각과의 차이가 0° 이상 112.5° 미만인면서, 타선의 선수각과 본선 간 사잇각, 그리고 타선 선수각과의 차이가 -112.5° 이상 10° 미만인 경우 :: 진로유지선 (그림 3 (b))

: 본선의 선수각과 타선 간 사잇각, 그리고 본선 선수각과의 차이가 -112.5° 이상 10° 미만인면서, 타선의 선수각과 본선 간 사잇각, 그리고 타선 선수각과의 차이가 0° 이상 112.5° 미만인 경우 :: 피항의무선 (그림 3 (c))

2. 심층 강화학습

본 연구에서는 워터젯의 출력 방향의 변화를 통해 CTV

를 제어하며, 워터젯의 출력 방향은 연속적인 데이터이다. 따라서 심층 강화학습에서는 연속적인 데이터인 워터젯의 출력 방향을 출력해야 한다.

기존 심층 강화학습 알고리즘 중 DQN (Deep Q Network)은 고차원의 입력 데이터 문제는 해결할 수 있지만, 저차원의 이산 출력 데이터만을 다룰 수 있다는 한계점이 존재한다 [13]. 또한, DDPG (Deep Deterministic Policy Gradient)는 학습 에이전트의 Q 함수가 Q 값을 극적으로 과대평가하여 정책 위반으로 이어지는 문제가 발생할 수 있다 [14].

TD3 (Twin Delayed Deep Deterministic Policy Gradient)는 연속적인 고차원의 출력 데이터도 다룰 수 있으며, 두 개의 Q 함수를 학습하고 그 중 작은 값을 사용하는 'Double critic network' 기법을 사용하여 Q 값을 극적으로 과대평가하여 발생할 수 있는 오류를 줄이고, Q 함수에 비해 더 큰 주기로 업데이트하여 인공신경망의 예측 정확도를 높일 수 있는 'Delayed policy updates' 기법을 사용한다 [14].

MATD3는 TD3를 다중 에이전트 환경에 적용한 알고리즘이다 [15]. 본 연구에서는 다중 에이전트를 제어할 수 있으며, 연속적인 고차원의 입력 데이터와 연속적인 고차원의 출력 데이터 처리에 좋은 성능을 보여주며, Q 값을 극적으로 과대평가하여 발생하는 오류를 줄이기 위해, MATD3를 사용해 에이전트의 학습 및 시험을 진행하였다.

MATD3는 그림 4와 같은 CTV의 장애물 회피 시스템 중 장애물 회피 기동 단계에서 선박의 환경과 관련된 변수를 입력하면 선박의 워터젯 출력 방향을 출력하여 장애물 회피 기동을 위해 사용되었다.

에이전트의 학습 및 시험에서 좋은 성능을 얻기 위해서는 입력 변수와 출력 변수의 형태를 설정하고 적절한 보상 함수를 설계해야 한다. MATD3의 입력 변수는 상태 공간을 의미하며, 출력 변수는 행동 공간을 의미한다.

2.1 상태 공간

상태 공간 (State space)은 MATD3 입력 변수의 집합으로 선박의 위치, 방향, 속도 등 다양한 변수가 사용될 수 있다. 총 n개의 선박이 존재하는 본 연구의 상태 공간 (S)은 각 선박의 위치 좌표 (x_i, y_i)와 선수각 (ψ_i)으로 이루어져 있다. 이를 식으로 표현하면 식 20과 같다.

$$S = [x_1, y_1, \psi_1, \dots, x_n, y_n, \psi_n]. \tag{20}$$

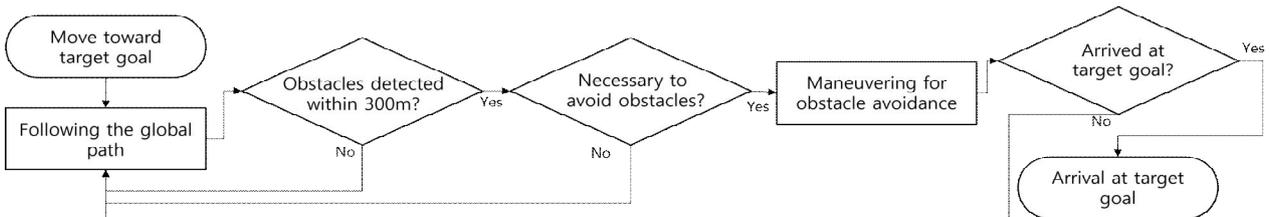


그림 4. CTV 운항 중 장애물 회피 시스템 흐름도
Fig. 4. Flowchart of obstacle avoidance system during CTV operation

2.2 행동 공간

행동 공간 (Action space)은 MATD3의 출력 변수의 집합이다. 본 연구에서는 워터젯 출력 시스템을 사용하고 워터젯의 출력은 고정되어 있다고 가정하며, 워터젯의 출력 방향만을 제어하여 장애물 회피 기동을 수행한다. 이에 따라 워터젯의 출력 방향을 인공신경망의 출력 변수로 선정하였다. 총 n 개의 선박이 있을 때 본 연구의 행동 공간 (A)은 식 21과 같이 표현할 수 있다.

$$A = [\delta_1, \delta_2, \dots, \delta_n]. \quad (21)$$

2.3 보상 함수

보상 함수 (Reward function)는 MATD3의 인공신경망이 장애물 회피를 위해 원하는 방향으로 어떻게 학습할 것인지를 설계할 수 있는 심층 강화학습의 중요한 요소이다. 만약 보상 함수 설계 단계에서 설계의 오류가 발생한다면 아무리 정확하고 많은 정보를 기반으로 학습을 수행해도 원하는 방향으로 학습되지 않아 잘못된 결과가 출력될 수 있다. 본 연구에서는 다음과 같이 보상 함수를 설계하였으며, 이를 식으로 표현하면 식 22과 같다.

- 장애물 회피를 위한 보상 함수
 - 장애물과의 거리에 대한 보상 (R_{OD})
 - 선박의 선수각과 장애물 간 사이 각에 대한 보상 (R_{OA})
 - 장애물과 충돌 여부에 대한 보상 (R_{OC})
- 목표지점 도착을 위한 보상 함수
 - 목표지점까지 거리에 대한 보상 (R_{GD})
 - 선박의 선수각과 목표지점 간 사잇각에 대한 보상 (R_{GA})
 - 목표지점 도착 여부에 대한 보상 (R_{GS})
- COLREGs 규정 준수를 위한 보상 함수
 - COLREGs 규정 준수 여부에 대한 보상 (R_{ξ})

$$R_{total} = R_{OD} + R_{OA} + R_{OC} + R_{GD} + R_{GA} + R_{GS} + R_{\xi}. \quad (22)$$

3. 시뮬레이션 학습 및 시험 환경 구성

MATD3의 큰 특징 중 하나는 배치 정규화를 통해 모델의 일반화 성능이 우수하다는 점이다. 이를 확인하기 위해 CTV의 장애물 회피 시뮬레이션 환경 구현 시 학습 환경과 시험 환경을 다르게 구성하였다.

3.1 MATD3 기반 장애물 회피 시뮬레이션 학습 환경

본 연구에서는 MATD3의 장애물 회피 학습을 위한 시뮬레이션 시나리오는 장애물의 종류에 따라 시나리오 이름을 다음과 같이 정의하였다.

- 본선과 전역 경로상 1개의 고정 장애물 (Fixed Obstacle, FO)이 존재 : '1 FO' - 그림 5 (a)
- 본선과 전역 경로상 1개의 이동 장애물 (Moving Obstacle, MO)이 존재 : '1 MO' - 그림 5 (b)

그림 5는 2가지 학습 시나리오의 개략도이며, 파란색 점은 본선의 시작점, 파란색 실선은 본선의 전역 경로, 노란색 별은 목적지, 검은색 점은 고정 장애물, 빨간색 실선은 타선의 전역 경로를 의미한다.

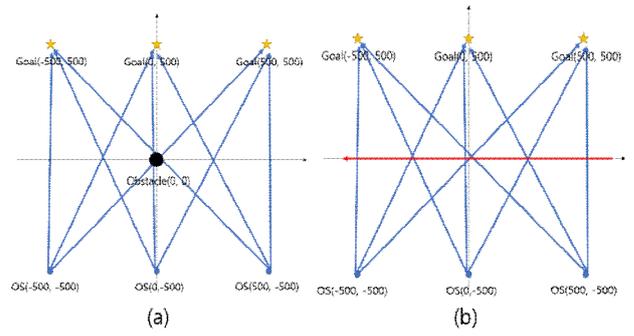


그림 5. 학습 시나리오 (a) '1 FO'와 (b) '1 MO'의 개략도
Fig. 5. Schematic of learning scenario (a) '1 FO' and (b) '1 MO'

표 2. 학습 시나리오 '1 FO'의 제원

Table 2. Specifications of learning scenario '1 FO'

Name	List	Value
Own Ship	Initial location	[-500, -500], [0, -500], [500, -500]
		Goal location
	Speed	10m/s
	Initial head angle	0°
Fixed Obstacle	Location	[0, 0]
	Diameter	20m

표 3. 학습 시나리오 '1 MO'의 제원

Table 3. Specifications of learning scenario '1 MO'

Name	List	Value
Own Ship	Initial location	[-500, -500], [0, -500], [500, -500]
		Goal location
	Speed	10m/s
	Initial head angle	0°
Moving Obstacle	Initial location	[500, 0]
	Goal location	[-500, 0]
	Speed	10m/s
	Initial head angle	270°

표 2는 '1 FO' 시나리오에서 사용한 초기 환경 변수를 나타낸 표이다. 해당 시나리오에서 최초 환경 변수는 본선의 최초 위치 및 목표 위치, 속도, 최초 선수각, 고정 장애물의 위치 및 크기로 구성되어 있다. 표 3은 '1 MO' 시나리오에서 사용한 초기 환경 변수를 나타낸 표이다. 해당 시나리오에서 최초 환경 변수는 본선과 이동 장애물인 타선의 최초 위치 및 목표 위치, 속도, 최초 선수각으로 구성되어 있다.

본 연구에서는 이와 같은 '1 FO' 시나리오와 '1 MO' 시나리오에 대해 각 1,000회, 5,000회, 10,000회의 에피소드만큼 학습을 수행하였다.

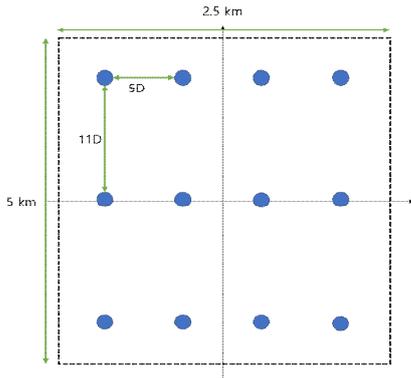


그림 6. 60MW 해상풍력발전단지 최적 설계 배치 개략도
Fig. 6. Schematic of optimal design layout plan for a 60MW offshore wind farm

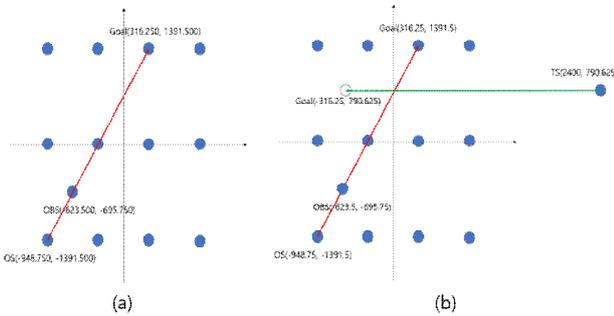


그림 7. 시험 시나리오 (a) '1 FO 0 MO'와 (b) '1 FO 1 MO'의 개략도
Fig. 7. Schematic of test scenario (a) '1 FO 0 MO' and (b) '1 FO 1 MO'

3.2 MATD3 기반 장애물 회피 시뮬레이션 시험 환경

MATD3의 장애물 회피 시험 시뮬레이션 환경은 이격거리에 따른 해상풍력발전단지 최적 배치에 관한 연구 [8]을 참고하여 블레이드 직경 (D)이 126.5m인 5MW급 해상풍력발전기 12개를 2.5Km * 4Km의 면적에 최적 설계하여 배치된 60MW급 해상풍력발전단지 환경을 가정하였다. 이와 동시에 45° 방향에서 불어오는 10m/s의 바람과 135° 방향에서 흐르는 조류가 존재하는 환경을 구현하여 장애물 회피 시뮬레이션 시험을 진행하였다. 5MW급 해상풍력발전기를 이용한 60MW 해상풍력발전단지 최적 배치는 그림 6과 같이 나타낼 수 있다. 그림에서 파란 점은 배치된 해상풍력발전기의 위치를 의미하며, 2.5Km * 4Km 구역 내에 가로로 5D만큼, 세로로 11D만큼 이격되어 배치된 모습을 확인할 수 있다.

시험 시나리오는 다음과 같은 2가지 시나리오로 나누어 시험하였으며, 45°도 방향에서 10m/s로 불어오는 바람과 135° 방향으로 1.7m/s의 속도로 흐르는 조류가 있는 상황으로 가정하였다. 그림 7은 2가지의 시험 시나리오의 개략도이다. 그림 7에서 파란색 점은 고정된 해상풍력발전기 또는 고정 장애물, 빨간색 실선은 본선의 전역 경로, 초록색 실선은 타선의 전역 경로를 의미한다.

- 본선과 해상풍력발전단지 외 고정 장애물 1개 : '1 FO 0 MO' - 그림 7 (a)와 표 4

표 4. 시험 시나리오 '1 FO 0 MO'의 제원
Table 4. Specifications of test scenario '1 FO 0 MO'

Name	List	Value
Own Ship	Initial location	[-948.75, -1391.50]
	Goal location	[-316.25, 1391.50]
Fixed Obstacle	Speed	10m/s
	Initial head angle	0°
	Location	[-623.50, -695.75]
	Diameter	40m

표 5. 시험 시나리오 '1 FO 1 MO'의 제원
Table 5. Specifications of test scenario '1 FO 1 MO'

Name	List	Value
Own Ship	Initial location	[-948.75, -1391.50]
	Goal location	[-316.25, 1391.50]
	Speed	10m/s
Fixed Obstacle	Initial head angle	0°
	Location	[0, 0]
	Diameter	40m
	Initial location	[500, 0]
Moving Obstacle	Goal location	[-500, 0]
	Speed	10m/s
	Initial head angle	270°

- 본선과 해상풍력발전단지 외 고정 장애물 1개와 이동 장애물 1개 : '1 FO 1 MO' - 그림 7 (b)와 표 5

3.3 장애물 회피 시뮬레이션 시험 결과 및 분석

시험 시나리오 '1 FO 0 MO'와 '1 FO 1 MO'는 MATD3의 입력 차원 문제로 각각 '1 FO' 학습 시나리오 모델과 '1 MO' 학습 시나리오 모델을 이용하여 시험을 진행하였다. 그 결과, 두 시험 시나리오 모두 1,000회 학습한 모델은 복잡한 해상풍력발전단지 시뮬레이션 환경에서 장애물을 회피하지 못하고 충돌하는 모습을 확인할 수 있었으며, 5,000회, 10,000회 학습한 모델은 바람과 조류의 영향이 있음에도 불구하고 장애물을 성공적으로 회피하고 목적지까지 도착한 결과를 확인할 수 있었다.

1,000회 학습한 모델이 장애물과 충돌한 이유는 각 1,000회, 5,000회, 10,000회 학습한 모델의 마지막 100회 동안의 평균 보상 점수를 기반으로 원인을 추측할 수 있다. 1,000회 학습한 모델의 평균 보상 점수는 5,000회, 10,000회 학습한 모델과 비교해 현저히 낮은 보상 점수를 획득하였으며, 이는 MATD3가 충분히 학습되지 않았음을 짐작할 수 있다. 따라서, 본 연구와 같은 환경에서 심층 강화학습에 기반하여 장애물을 회피 기동하기 위해서는 최소 5,000회의 반복 시뮬레이션을 통해 충분한 학습이 이루어져야 한다는 것을 알 수 있다. 그림 8은 '1 FO' 학습 시나리오를 10,000회 학습한 모델을 이용해 시험 시나리오 '1 FO 0 MO'에서 시험한 결과, 그림 9는 '1 MO' 학습 시나리오를 10,000회 학습한 모델을 이용해 시험 시나리오 '1 FO 1 MO'에서 시험한 결과이다. 각 그림에서 검은색 원과 원의 중심은 선박의 장애물 인식 거리인 300m 반경 및 선박의 위치를 의미하며, 검

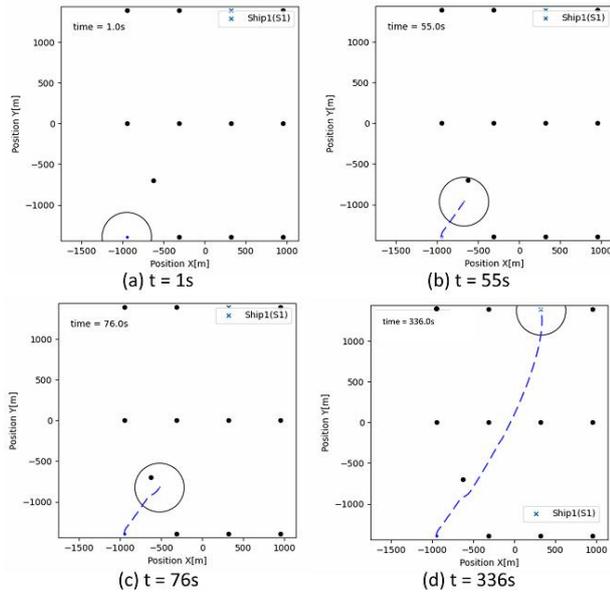


그림 8. 시험 시나리오 '1 FO 0 MO'의 결과
Fig. 8. Result of test scenario '1 FO 0 MO'

은색 점은 고정 장애물, 파란색 점선은 본선의 이동 경로, 빨간색 점선은 타선의 이동 경로를 의미한다. 파란색 x와 빨간색 x는 각각 본선과 타선의 목적지를 의미한다.

시험 시나리오 '1 FO 0 MO'인 그림 8에서는 (b) $t=55s$ 일 때, 선박은 장애물과 거리가 300m 이하가 된다. 이때, 선박은 고정 장애물과의 충돌 위험이 판단되면 심층 강화학습 네트워크 모델은 (c) $t=76s$ 와 같이 워터젯의 출력 방향을 제어하여 장애물을 회피한다. 장애물 회피 후 (d) $t=336s$ 인 마지막까지 충돌하지 않고 목적지까지 도달하는 모습을 확인할 수 있다.

시험 시나리오 '1 FO 1 MO'인 그림 9에서는 시험 시나리오 '1 FO 0 MO'와 같이 첫 번째 고정 장애물을 회피하고 $t=232s$ 일 때, 이동 장애물과 거리가 300m 이하가 된다. 이때 선박은 선박의 우현 방향으로부터 다가오는 이동 장애물과 충돌할 위험이 있다고 판단하고 심층 강화학습 네트워크 모델은 $t=242s$ 와 같이 워터젯의 출력 방향을 제어하여 장애물을 회피한다. 장애물 회피 후 (d) $t=336s$ 인 마지막까지 어떠한 고정 장애물 및 이동 장애물과 충돌하지 않고 목적지까지 도달하는 모습을 확인할 수 있다.

IV. 결론

본 연구에서는 해상풍력발전단지 내에서 해상풍력발전기의 유지보수를 위해 운영되는 CTV의 장애물 회피에 관해 연구하였다. 특히, COLREGs를 기반으로 심층 강화학습을 이용한 CTV의 장애물 회피에 관해 연구하였다.

이를 위해 연구에서는 CTV의 조종 운동 모델링을 수행하고, 해상풍력발전단지과 유사한 시뮬레이션 환경 구현을 위해 실제 관측데이터를 기반으로 풍하중과 조류 하중을 구

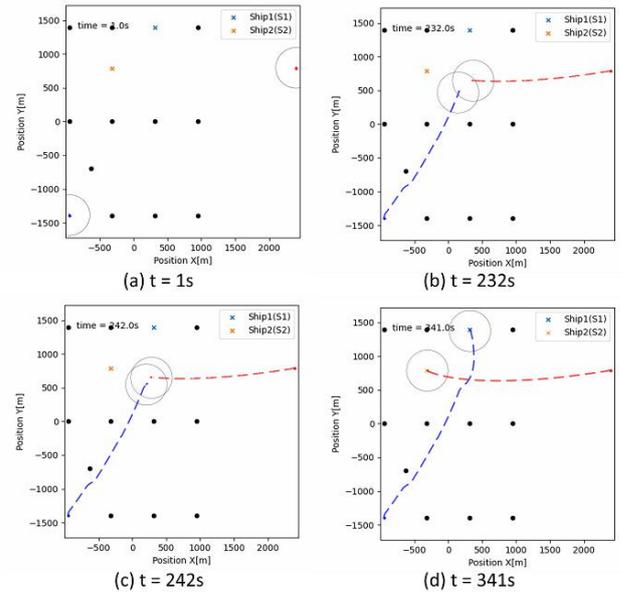


그림 9. 시험 시나리오 '1 FO 1 MO'의 결과
Fig. 9. Result of test scenario '1 FO 1 MO'

현하였다. 또한 COLREGs를 고려해 고정 장애물 및 이동 장애물과의 충돌 여부를 판단하고, 이를 기반으로 장애물 회피 기동 필요 여부를 판단하였다. 장애물 회피 기동을 위한 워터젯의 출력 제어는 심층 강화학습 중 고차원 입력 데이터와 고차원 출력 데이터를 다룰 수 있으며 다중 에이전트 제어 시 성능이 좋은 MATD3를 사용하였다. 시뮬레이션 환경은 해상풍력발전단지의 최적 배치를 참고하여 학습 및 시험 시뮬레이션 환경을 구현하였다.

심층 강화학습을 이용한 장애물 회피 시험 시뮬레이션 결과, 학습 단계에서 충분한 학습이 이루어진다면 시뮬레이션 상에서 CTV가 장애물과 조우 시 성공적으로 장애물을 회피할 수 있음을 확인할 수 있었다.

본 연구에서는 실제 CTV 선박이 아닌 무인수상선을 대상으로 조종 운동 모델링을 수행하였으며, 시뮬레이션 환경에서 장애물 인식을 위한 장애물 탐지 시스템을 구현하지 않았다. 또한, 일정하게 유지되는 외부 환경의 시뮬레이션 환경에서 연구를 수행하였다. 따라서, 향후 실제 CTV의 조종 운동 모델링, 장애물 탐지 시스템, 시간에 따라 랜덤하게 변하는 환경 하중을 고려한 시뮬레이션 환경에 관한 연구가 수행될 필요가 있다.

References

- [1] J. N. Kim, "Carbon-neutral Technological Innovation Promotion Strategy - 10 core Technology Development Directions," Korea Institute of Energy Research, 2021 (in Korean).
- [2] H. K. Cho, B. S. Park, D. H. Kang, S. S. Kim, "The Main Dactor and Counterplan for Marine Accidents in Korea,"

Journal of Fisheries and Marine Sciences Education, Vol. 29, No. 7, pp. 746-756, 2017 (in Korean).

[3] D. G. Kim, J. S. Jeong, G. K. Park, "A Study on Ship Collision Avoidance Algorithm by COLREG," *International Journal of Fuzzy Logic and Intelligent Systems*, Vol. 21, No. 3, pp. 290-295, 2011 (in Korean).

[4] D. G. Son, D. H. Kim, "Autonomous Navigation of Unmanned Surface Vehicle Based on Fuzzy Control Considering COLREG," *Journal of Korean Institute of Intelligent Systems*, Vol. 31, No. 1, pp. 11-20, 2021 (in Korean).

[5] J. H. Woo, N. Kim, "Collision Avoidance for an Unmanned Surface Vehicle Using Deep Reinforcement Learning," Doctor's thesis, Seoul National University, 2018 (in Korean).

[6] N. Wen, Y. Long, R. Zhang, G. Liu, W. Wan, D. Jiao, "COLREGs-based Path Planning for USVs Using the Deep Reinforcement Learning Strategy," *Journal of Marine Science and Engineering* 11, No. 12, pp. 23-34, 2023.

[7] X. Xu, P. Cai, Z. Ahmed, V. S. Yellapu, W. Zhang, "Path Planning and Dynamic Collision Avoidance Algorithm Under COLREGs Via Deep Reinforcement Learning," *Neurocomputing* 468, pp. 181-197, 2022.

[8] Y. J. Kim, "A Study on Optimal Design of Offshore Wind Farm Based on Array Matrix," Master's thesis, Mokpo National University, 2013 (in Korean).

[9] J. S. Jeong, "Design of an Auto-pilot for the Towing Tank Test of Model Ship," Master's Thesis, Chungnam National University, 2005 (in Korean).

[10] T. I. Fossen, "Guidance and Control of Ocean Vehicles," John Wiley & Sons Inc, 1994.

[11] H. Qu, E.I. Sarda, I. R. Bertaska, K. D. Ellenrieder, "Wind Feed-forward Control of a USV," *IEEE*, pp. 1-10, 2015.

[12] W. Y. Hwang, "Application of System Identification to Ship Maneuvering," PhD thesis, Massachusetts Institute of Technology, 2022.

[13] T. P. Lillicrap, J. J. Hunt, A. Pritzel, T. Erez, Y. Tassa, D. Silver, D. Wierstra, "Continuous Control with Deep Reinforcement Learning," *arXiv preprint arXiv:1509.02971*, 2016.

[14] S. Fujimoto, H. Hoof, D. Meger, "Addressing Function Approximation Error in Actor-critic Methods," *PMLR*, 80:1587-1596, 2018.

[15] M. Zhan, J. Chen, C. Du, Y. Duan, "Twin Delayed Multi Agent Deep Deterministic Policy Gradient," *IEEE*, pp. 48-52, 2021.

Jingyun Kim (김진균)



2019 Naval Architecture and Ocean Engineering from Kunsan National University (B.S.)
 2023 Naval Architecture and Ocean Engineering from Kunsan National University (M.S.)
 2023~Naval Architecture and Ocean Engineering from Kunsan National University (Ph.D. cadidate.)
 Field of Interests: Deep Reinforcement Learning Application
 Email: wlsrbs0303@kunsan.ac.kr

Jaemyung Chon (전해명)



2019 Naval Architecture and Ocean Engineering from Kunsan National University (B.S.)
 2020 Naval Architecture and Ocean Engineering from Kunsan National University (M.S.)
 2022~Naval Architecture and Ocean Engineering from Kunsan National University (Ph.D. cadidate.)
 Field of Interests: Deep Learning Application
 Email: tipptipp@kunsan.ac.kr

Jackyou Noh (노재규)



1996 Naval Architecture and Ocean Engineering from Seoul National Univerisy (B.S.)
 1998 Naval Architecture and Ocean Engineering from Seoul National Univerisy (M. S.)
 2009 Naval Architecture and Ocean Engineering from Seoul National University, Seoul, Republic of Korea (Ph.D.)
 2010~Naval Architecture and Ocean Engineering, in Kunsan National University (Prof.)
 2014~Korean J. of Computational Design and Engineering, Associate Editor
 Field of Interests: Systems Control, and Deep Learning
 Email: snucurl@kunsan.ac.kr