AI-based Collision Avoidance for Automatic Ship Navigation

Hitoshi Yoshioka, *Graduate School of Engineering, Osaka Prefecture University*
sbb03124@st.osakafu-u.ac.jp
Hirotada Hashimoto, *Graduate School of Engineering, Osaka Metropolitan University*
hashimoto.marine@omu.ac.jp

ABSTRACT

In this research, AI-based collision avoidance is developed for autonomous ship navigation. The danger of collision is evaluated using Dangerous Area of Collision (DAC). The DAC and a waypoint are given as state for the agent in a value-mapping style in which a value corresponding to a condition is assigned to each grid of domain. Negative rewards are given in a reinforcement learning if other ships enter the DAC. As a result, the developed AI can navigate to the given waypoint and avoids collision if necessary. A numerical experiment is conducted for some congested situations using actual measurement data at sea, and it is demonstrated that AI-based collision avoidance can avoid the possible collisions effectively. Through the numerical validation, it is concluded the AI-based autonomous navigation can be achieved with a reasonable safety margin.

Keywords: *AI-based collision avoidance, Deep Q-learning, Dangerous area of collision, Automatic navigation.*

1. INTRODUCTION

Collision avoidance is always a crucial issue for ship safety as it depends on the judgment and actions of seafarers. Most of ship collisions are caused by human errors. As long as humans operate a ship in a conventional way, it is essentially hard to prevent collisions perfectly. Therefore, a machine-based collision avoidance as a navigation supporting system is expected to prevent ship collisions own to human errors, for realizing safer navigation. In addition, automatic collision avoidance is an essential function for autonomous ships in the future. In congested waters, there are many collision risks with surrounding ships. Further, it is necessary to make an appropriate decision for collision avoidance from not only collision risks but also waypoints, external disturbance, shoals, and so on. It is difficult to explicitly model a decision-making process in collision avoidance by veteran captains in such complicated situations. Recently, it is expected to realize advanced collision avoidance which can handle such complex process by using AI technology. Many studies on autonomous collision avoidance are reported as AI technology is rapidly developing.

Ship collision avoidance algorithms based on deep reinforcement learning were developed and they were validated by a free-running model experiment (Shen et al., 2019) or by a numerical experiment (Sawada et al., 2021). These studies show that AI-technology is effective for collision avoidance in various encountering situations. The difference between a simulation and an actual ship experiment was reported recently (Hashimoto et al., 2021). In previous research, the AI was developed by deep Q-learning using the detection lines and the predicted area of danger (PAD) (Bole et al., 2005) to describe the state for the neural network input. In the study, the number of detection lines are limited and hence the possible danger of collision cannot be fully detected. In addition, the shape of the PAD was simplified to ease the calculation for detection. Although these matters were not so influential for the problems discussed in the previous research, i.e., collision avoidance in congested and confined waters, further efforts might be necessary for practical uses in an advanced navigation supporting system.

In this study, an autonomous collision avoidance algorithm is presented which is enhancing the existing algorithm (Shen et al., 2019). The input data construction is newly designed using a grid and value assignment. The Dangerous Area of Collision (DAC) (Hakoyama et al., 1996) was used to illustrate the collision risk area. In addition, a fundamental function to sail to the given waypoint is achieved using the same grid as for collision avoidance. The developed AI is validated for real encountering situations obtained by the past actual ship experiment to demonstrate its effectiveness as for the autonomous navigation in the future.

2. LEARNING METHODS

Deep Q-learning

Deep Q-learning (Mnihm et al., 2013) is one of methods of deep reinforcement learning and is applied in many fields. In reinforcement learning, there are an agent, a state, and a reward in environment. According to the agent's action taken, the state transits to a next state and then, the agent gets new observation of state and reward. The agent decides an optimal action from the observation. For this purpose, the agent learns the action which maximizes a cumulative reward in future. The cumulative value of reward in future is called Q-value. In the deep Q-learning, an action-value function expressed by a multi-layer neural network (NN), so-called deep Q-network, is used. The input for NN is a state which the agent gets from environment and the output from NN is Q-values for actions. In learning process, parameters of neural network are optimized to minimize a loss function. Here, the optimal action means the action which is expected to gain the maximum Q-value among actions.

Manoeuvring Model

A ship manoeuvring motion is calculated by Nomoto's K-T model (Nomoto, 1960) for the learning, which is shown in equation (1). The model is expressed as equation (1) and used for learning:

$$T\dot{r} + r = K\delta \tag{1}$$

where r, δ, T, K are rate of turn, rudder angle, time coefficient, and gain, respectively. The values of K and T are 0.183 1/s and 11.1 s.

Dangerous Area of Collision (DAC)

To illustrate the collision risk area, the so-called Dangerous Area of Collision (DAC) (Hakoyama et al., 1996) was applied. DAC is one of methods to display dangerous area. Safe Passing Area (SPA) around the own ship is defined as a circle. Its radius means a minimum value as the safe distance between the own ship and a target ship. In the original DAC calculation method, SPA is approximated by a polygon and virtual own ships are placed at vertexes, denoted as $a, b, \dots, f$ in Fig.1. The points where virtual own ships collide with the target ship, denoted as $a', b', \dots, f'$, are calculated. The places of own ship when virtual ships collide with the target ship, denoted as $a'', b'', \dots, f''$, are determined from the collision points and relative positions between own ship and vertexes. DAC is defined as the area drawn by connecting the projected points. The projection of each vertex and DAC are shown in Fig.1. The details of calculation procedure and an application example can be found in literature (Hashimoto et al., 2022).

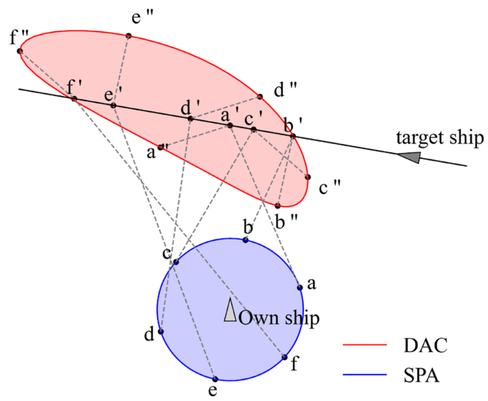


Fig.1 Image of vertex projection in DAC calculation

Construction of state

The input for NN is a state and it should include information required for decision-making of collision avoidance. For the construction of state for collision avoidance, a gray scale image with a grid system is used. This might be similar information to ECDIS and/or radar screen. The domain is a square area around the own ship with length and width of 14 km. The size of domain is set according to the maximum range of an evaluation area diagram for collision avoidance manoeuvring (Nakamura and Okada, 2019). The domain is divided into finite sections, (42×42) in this study. The length of dangerous area in front of a ship is generally longer than that on behind. Therefore, the own ship is placed in the domain with 3.5 km off-set to the front. As a result, the fore length is 10.5 km and the aft length is 3.5 km from the own ship to the end of domain. For the transversal direction, the same length (7.0 km) is used for both sides.

In order to evaluate the level of collision danger, the radius of SPA should be appropriately selected because the danger of collision increases as the radius of SPA decreases. The change of danger levels is illustrated in Fig.2 which shows DACs with different radius of SPA. Fujii (1980) proposed a shape of ship domain and it is described with ellipse for fore and circle for aft. The lengths of major and minor axes are $6.4L$ and $1.6L$, where L means length overall of the ship. Because the shape of SPA for the DAC is circle, the radius is determined as $3.2L$ to have the same value of product. The danger is set to 10 levels and each radius is determined by equally dividing the maximum radius. The SPA with radius of $3.2L * i/10$ is denoted as SPA_i. The DAC calculated by projecting SPA_i is denoted as DAC_i. The level of danger increases when the distance to other ships decreases, so the higher value is given to closer area as shown in Fig.2.

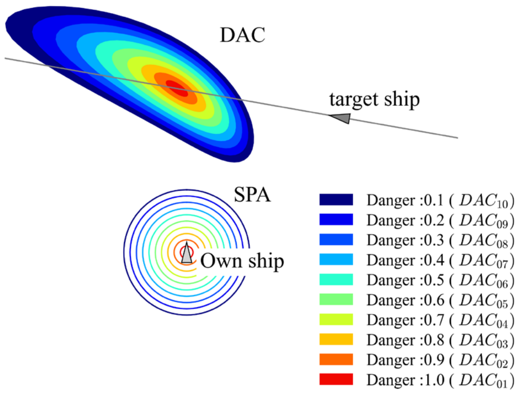


Fig.2 Level of collision illustrated by DACs with different radius of SPA

For autonomous navigation, it is necessary to automatically sail to given waypoints, to automatically avoid collisions with reasonable safety margin, and to automatically recover to the original course after the collision avoidance. For this purpose, information of a waypoint is included in input data within the same framework of collision avoidance. The value between 1.0 and -1.0 is given to all cells consisting of the domain except for cells occupied by the DAC. The value for each cell is calculated by the distance to the waypoint. When the distance between a cell and the own ship is zero, -1.0 is given. When the distance is equal to or longer than the distance between the waypoint and the own ship, 0.0 is given. The value is changed linearly with the distance. In addition, -0.1 is given when a cell contains a predicted future trajectory of the own ship. The future trajectory is predicted for 3 minutes. The own ship crosses 4 cells for 3 minutes at the service speed so that it can be judged whether the own ship is turning or not. An example of input data for NN is shown in Fig.3.

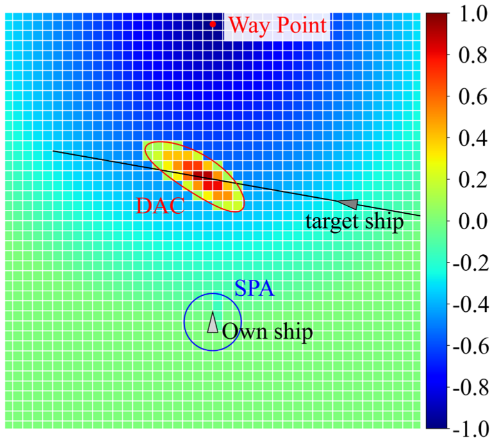


Fig.3 An example of input data to NN

Reward setting

Rewards for agent in reinforcement learning are set as shown in Table 1. In learning process, the parameters of NN are optimized to predict expected total cumulative reward. To realize preferable collision avoidance manoeuvre according to COLREGs, a penalty for left turning is set greater than right turning.

Table 1 Reward setting

	reward
sailing to waypoint	0 to 0.1
overlapping of SPA and DAC	0 to -1.0
overlapping of SPA and other ships	-1.5
left turning	-0.25
right turning	-0.05

Action

In deep Q-learning, the agent sequentially selects an action which is expected to gain the maximum Q-value at each timestep. In this study, a discretized rudder angle is used as an action. The action options are shown in Table 2. The actions are three rudder angles to keep its course or turn to left or right, and the change of speed is not allowed for the simplicity. A ship manoeuvring motion induced by a selected action, i.e., rudder angle, is calculated with the Nomoto's K-T model mentioned before at a timestep of 10 seconds.

Table 2 Agent's actions

purpose	action (rudder angle)
turn right	5[deg]
sail straight	0[deg]
turn left	-5[deg]

Other ship

In the learning environment, the number of other ships is decided randomly. A heading angle and speed of other ships and the waypoint are also decided randomly. The range of each parameter is shown in Table 3. Other ships do not change their course and sail straight only.

number of ships	0 ~ 30
initial position	random place in 14 km squared area
heading angle	0 ~ 360 [deg]
service speed	5 ~ 10 [knots]
initial waypoint	60 miles from own ship

Neural network

The neural network structure used for study is shown in Table 2. The convolutional layers (Conv) learn features including spatial information. Kernel size, slide amount, and number of layers are hyperparameters and were decided using convolutional autoencoder. Convolutional autoencoder has encoding and decoding functions. If the decoder can decode output data of the encoder to original data, it is meant the output data of encoder contains sufficient features. Several encoders were tested, and the structure of encoder resulting in the best performance was adopted for Conv.

Table 2 Structure of neural network

Input layer			
Conv	filters	kernel size	strides
	32	5	3
Conv	filters	kernel size	strides
	64	3	1
FC	nodes: 512		
FC	nodes: 128		
FC	node: 3		
Output layer			

Fully connected layers (FC) give the expected cumulative reward for each possible action. The numbers of layers and cells of each layer are important hyper parameters. If the number of nodes is too large for the problem, the value of weights of NN doesn't converge. On the other hand, the performance of NN becomes poor if there are not enough nodes. The number of layers and nodes of FC were determined by trial and error.

3. EVALUATION AND DISCUSSIONS

Evaluation method

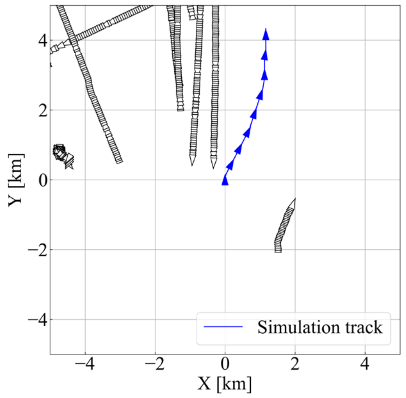
A subjective evaluation method for collision avoidance manoeuvres is used to evaluate the AI. This method was developed by analysing a lot of results of simulator experiment in collision avoidance by captains. Collision risk is evaluated by distance between the own ship and a target ship and changing rate of relative bearing to a target ship. The details can be found in the literature (Nakamura and Okada, 2019).

Evaluation results

The AI manoeuvre is evaluated for several encountering situations, logged by onboard sensors in a past actual ship experiment (Hashimoto et al., 2021). The same manoeuvring model as for the learning is used for the evaluation. The own ship is controlled by the developed AI and other ships sail as same as the observed data. The time step for the simulation is 10 seconds.

The results are shown in Figs.4-6. The upper and lower figures show the trajectory of the AI manoeuvre and evaluation result, respectively. The waypoint for all scenarios is set on the far forward direction. In the figure, a blue line and triangles show the trajectory and heading angle of the own ship. Other ships' position and heading angle are also plotted in black colour. In the evaluation result, "white zone", "yellow zone" and "red zone" are categorized as "safe zone", "caution zone", and "danger zone". The absolute distance and changing rate of relative bearing are calculated for all other ships every timesteps to evaluate the AI's manoeuvring. The manoeuvring result is evaluated with the frequency of invasion into danger and caution zones.

Although the tested situations are congested as a whole, the developed AI can navigate to the waypoint and avoid collisions successfully. It is confirmed that the ship is heading to the waypoint again after the safety is secured. In all cases, invasion into the danger zone is not observed and only few invasions into the caution zone are observed. This means that AI can avoid collisions with suitable safety margin for congested encountering situations.



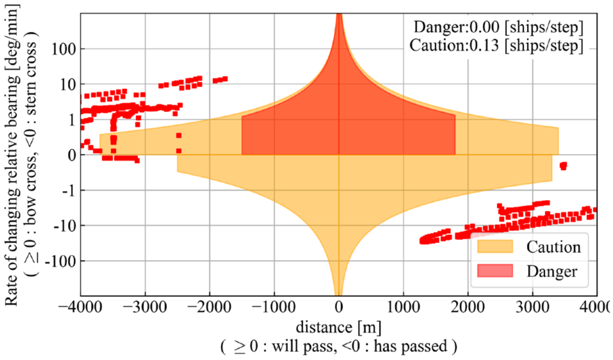
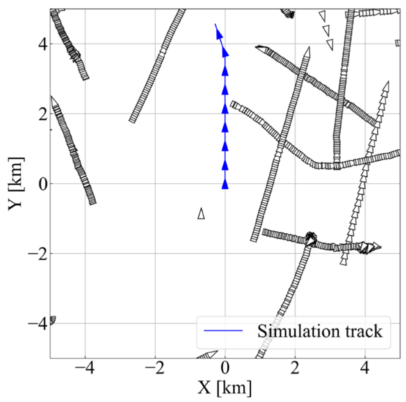


Fig.4 Ship trajectory and evaluation result of AI maneuver (Case 1)



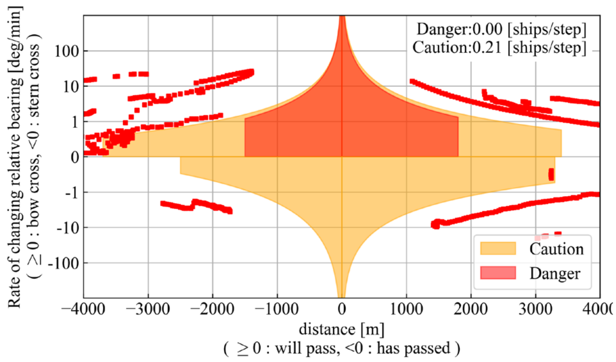
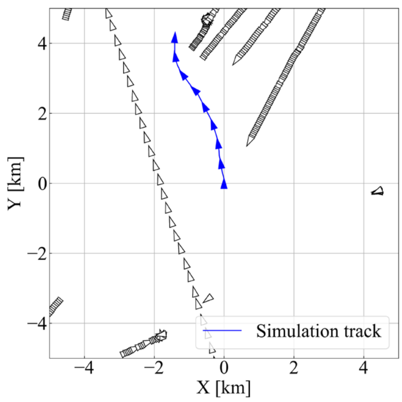


Fig.5 Ship trajectory and evaluation result of AI maneuver (Case 2)



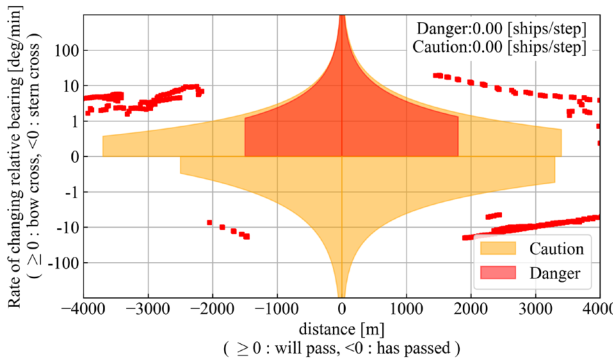


Fig.6 Ship trajectory and evaluation result of AI maneuver (Case 3)

4. CONCLUSIONS

AI for collision avoidance is developed by deep Q-learning for autonomous ship navigation. In the reinforcement learning, the danger of collision area/level is evaluated using Dangerous Area of Collision (DAC). The DAC and a waypoint are given as input for a multi-layer neural network in the same way, in which normalized value corresponding to the situation is assign to each cell consisting of the domain. Negative rewards are given when other ships or DAC enter the Safe Passing Area (SPA).

Through the numerical experiment for real encountering situations, it is demonstrated the AI-based manoeuvre can navigate to the given waypoint, avoid collision with reasonable safety margin, and return to the original course. It implies that AI can navigate in situations where human-operated vessels exist.

Toward realizing autonomous navigation in the future, further advancement in algorithms and quantitative validation as well as actual ship experiment for validation are expected.

REFERENCES

Bole, A., Dineley, B. and Wall, A. Radar and Arpa Manual, Elsevier, 2005

Fujii, Y., A definition of the evasive domain. (In Japanese), The journal of Japan Institute of Navigation, Vol.65, pp.17–22, 1980.

Hakoyama, T., Kato, Y., Maeda, S., Yamaguchi, H. and Yagi, O., Navigation display apparatus for collision avoidance utilizing polygonal safety regions and predicted danger areas. United States Patent." Patent Number: US5515287A, Date of Patent: May 7, 1996.

Hashimoto, H., Nishimura, H., Nishiyama, H. and Higuchi, G., Development of AI-based Automatic Collision Avoidance System and Evaluation by Actual Ship Experiment, ClassNK Technical Journal, No.3, 2021.

Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D. and Riedmiller, M., Playing Atari with Deep Reinforcement Learning. Tech. report. Deep. Technol. arXiv1312.5602 [cs.LG], 2013.

Nakamura, S. and Okada, N., Development of Automatic Collision Avoidance System and Quantitative Evaluation of the Manoeuvring Results, the International Journal on Marine Navigation and Safety of Sea Transportation (TRANSNAV), Vol.13, No.1, pp.133-141, 2019.

Nomoto, K., Analysis of Kempfs standard maneuver test and proposed steering quality indices, Proceedings of the 1st symposium on ship manoeuvrability. pp 275–304, 1960.

Sawada, R., Sato, K. and Majima, T., Automatic ship collision avoidance using deep reinforcement learning with LSTM in continuous action spaces, Journal of Marine Science and Technology, Vol.26(2), pp.509-524, 2021.Shen, H., Hashimoto, H., Matsuda, A., Taniguchi, Y., Terada, D. and Guo, C., Automatic collision avoidance of multiple ships based on deep Q-learning, Applied Ocean Research, Vol.86, pp.268-288, 2019.

Hashimoto, H., Makino, H., Yoshioka, H. and Matsuda, A., Ship-stopping algorithm utilizing VecTwin rudder system for automatic collision prevention, Ocean Engineering, Vol.251, 111098, 2022.