Journal of Automation and Intelligence 1 (2022) 100003

Contents lists available at ScienceDirect

Journal of Automation and Intelligence

journal homepage: www.keaipublishing.com/en/journals/journal-of-automation-and-intelligence/

A review on COLREGs-compliant navigation of autonomous surface vehicles: From traditional to learning-based approaches☆

Liang Hu [a,*], Huosheng Hu [a], Wasif Naeem [b], Zidong Wang [c]

[a] School of Computer Science and Electronic Engineering, University of Essex, Colchester, CO4 3SQ, UK
[b] School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast BT7 1NN, UK
[c] Department of Computer Science, Brunel University London, Uxbridge UB8 3PH, UK

ARTICLE INFO

Keywords:
Autonomous surface vehicle
Collision avoidance
Path re-planning
Deep reinforcement learning

ABSTRACT

A growing interest in developing autonomous surface vehicles (ASVs) has been witnessed during the past two decades, including COLREGs-compliant navigation to ensure safe autonomy of ASVs operating in complex waterways. This paper reviews the recent progress in COLREGs-compliant navigation of ASVs from traditional to learning-based approaches. It features a holistic viewpoint of ASV safe navigation, namely from collision detection to decision making and then to path replanning. The existing methods in all these three stages are classified according to various criteria. An in-time overview of the recently-developed learning-based methods in motion prediction and path replanning is provided, with a discussion on ASV navigation scenarios and tasks where learning-based methods may be needed. Finally, more general challenges and future directions of ASV navigation are highlighted.

1. Introduction

Research and development of autonomous surface vehicles (ASVs) have been conducted in both academia and industry worldwide for a number of years [1–4]. One basic requirement for ASVs is that they should be able to navigate autonomously and avoid collisions with other ships or obstacles. Furthermore, ASVs should behave and operate in a manner similar to that of manually operated vessels. Therefore, ASVs should be equipped with intelligent collision avoidance and motion planning techniques for their safe operations in dynamic and complex waterways.

Research in ASV navigation systems started in the early 2000s [5–9]. All manned ships are required to follow the rules of COLREGs (Convention on the International Regulations for Preventing Collisions at Sea) defined by the International Maritime Organisation (IMO) [10, 11], so it is imperative to impose COLREGs-compliant behaviours as an integral element of ASV navigation systems. The COLREGs rules are written for human consumption and are subject to various interpretations due to their subjective nature. As a result, how to ensure ASVs' behaviours in accordance with COLREGs has remained a central topic in the research of ASV navigation.

In this review, we overview the development on ASV COLREGs-compliant navigation from two dimensions: the complexity of ship encounter scenarios and the evolution in methodologies, as depicted in Fig. 1. As to ship encounter scenarios, early research started with COLREGs-compliant navigation in simple, 1-1 ship encounter scenarios, later research further considered multi-ship encounter scenarios even involved with COLREGs non-compliant vessels, and most recent research targeted at more complex scenarios such as areas of restricted visibility and busy narrow channels with traffic separation scheme where good seamanship from experienced mariners is needed. The progress in ship encounter scenarios is also observed in test and validation. Most early research was validated only by computer simulation scenarios. Later on, as lab-scale ASVs were built by some research labs, carefully designed on-water scenarios were used to test ASV navigation systems [12–15]. Very recently, quite a few real-ship tests using ASVs have been conducted at sea directly where real-world marine traffic scenarios are involved [16–19].

As to the methods for ASV navigation, traditional methods mainly come from (manned) ship navigation and robotics research. Many risk assessment and decision making techniques developed for (manned) ship navigation have been borrowed for ASV navigation. Furthermore, various robot collision avoidance and real-time motion planning techniques have been adapted for autonomous COLREGs-compliant marine navigation. However, it is hard to embody good seamanship, the highest requirement in COLREGs, in ASV navigation systems developed using traditional methods. Being COLREGs-compliant does not mean an ideal evasive behaviour, as the path might be over-conservative or inattentive to unexpected behaviour from target ships (TSs). Thanks to

☆ This work was supported in part by the Engineering and Physical Sciences Research Council (EPSRC) of the U.K., the Royal Society of the U.K.
* Corresponding author.
E-mail addresses: l.hu@essex.ac.uk (L. Hu), hhu@essex.ac.uk (H. Hu), w.naeem@ee.qub.ac.uk (W. Naeem), zidong.wang@brunel.ac.uk (Z. Wang).

https://doi.org/10.1016/j.jai.2022.100003
Received 16 June 2022; Received in revised form 23 October 2022; Accepted 3 November 2022

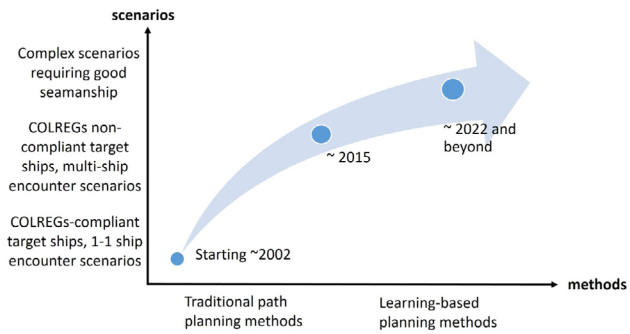


Fig. 1. The technical trend of ASV navigation systems in the past two decades.

the availability of huge amount of automatic identification system (AIS) data [20] and powerful deep machine learning algorithms especially deep reinforcement learning (DRL), promising solutions that ensure ASV navigation not only complies COLREGs rules but also imitates good seamanship of experienced mariners have been obtained. The development of both traditional methods and newly emerging learning based methods are reviewed in this paper.

To limit the scope of this review, we focus on the collision avoidance system (CAS) that enables COLREGs-compliant navigation and summarise the techniques and methods developed for three different subsystems of the CAS: risk detection, decision making and motion planning. Other broad research fields of ASVs such as sensing and localisation, guidance and control are omitted and instead the reader is referred to recent comprehensive surveys on the subjects [2,20].

In literature, different terms such as ASVs, unmanned surface vehicles (USVs), autonomous ships/boats/craft are commonly used as synonyms. Even though slight variations may exist among these terms for CASs, very little difference exists among these terms. As such, for simplicity and to keep the discussion generic, we use the terminology, ASVs in this paper to refer to all the above-mentioned terms.

1.1. Related literature reviews and our contribution

In the past decade, a number of review papers have been published that summarise research progress in navigation and motion planning of ASVs. The most relevant review papers are listed below:

- [1] provides the first published review on development of collision avoidance methods to increase autonomy of ASVs. The review was published more than one decade ago and hence did not cover recent research progress in ASV navigation.
- [2] reports a comprehensive review of ASV development from all aspects of guidance, navigation and control. Since only one section of the entire paper is focused on collision avoidance and path re-planning techniques, a brief summary and discussion is provided.
- [21] presents a comprehensive review on ship collision avoidance techniques of both manned and unmanned ships, with in-depth comparison and analysis of the pros an cons of different methods.
- [3] provides an in-depth review of traditional collision avoidance and path planning techniques of ASVs, without a review of the newly developed learning-based methods.

Next, we will state the intention and contribution of our review paper and highlight the need of this work compared with the existing review papers:

- We provide an up-to-date review of ASV navigation techniques. Not only the traditional methods, but also recent developments in learning based ASV navigation techniques are also covered, which is rarely discussed in detail in previous review papers. Particularly, there has been a surge in research interest in developing DRL-based collision avoidance and path re-planning methods for ASVs since the mid-2010s, however, the DRL-based approach for ASV navigation has not yet been reviewed systematically by any review papers so far. We would like to provide an in-time review depicting the research trends of ASV navigation in this paper.
- We provide a comprehensive review of collision avoidance strategies of ASVs, examining technical developments in all three subsystems of ASV CASs: risk detection, decision making and motion planning. With a holistic viewpoint, we could pinpoint and discuss clearly the challenges and possible solutions for safe navigation of ASVs in the future.

This review is divided as follows: Section 2 introduces the autonomous navigation systems of ASVs and CORELGs rules; Section 3 provides a comprehensive survey of ASVs models, collision detection, decision making and path replanning; Section 4 reviews the recently-developed DRL techniques and the DRL-based ASV navigation. Finally, in Section 5, the concluding remarks are given.

2. Autonomous navigation systems

In this section, we first briefly introduce the navigation system of ASVs, then summarise the maritime traffic rules on the sea, i.e., the COLREGs, and finally discuss the challenges faced by ASVs operating in the city's waterways. In this review, we assume that the ASV equipped with the CAS is the own ship (OS) whilst all other vessels around the OS are referred to as TSs, unless otherwise stated.

2.1. System architecture

The navigation system of an ASV is decomposed into four functional modules, perception and localisation, global path planning, collision avoidance and vessel control, as depicted in Fig. 2. The perception and localisation system provides situational awareness of the ASV and its environment, such as the location and motion information of the ASV and other moving vessels detected by the ASV's on-board sensors. The route planner of the ASV plans a global path (route) to the destination through traversable water areas, which is followed by a CAS whose task is to re-plan the path locally in case of collision risks with other vessels in the neighbouring area. A vessel controller adjusts the ASV's speed and heading such that the ASV follows the planned path smoothly.

2.1.1. Perception and localisation

Marine craft are usually equipped with a combination of different sensors including radar, lidar, camera, sonar, GPS and inertial measurement units, to name a few. The measurement from different sensors are fused together using signal processing and state estimation algorithms, producing accurate estimate of the OS and TSs' states such as position, orientation, speed and heading. Additionally, good understanding of the surrounding environments is also obtained, such as the relative location and size of obstacles, the current and tide speeds.

2.1.2. Route planning

Route planning, which is also called global path planning, is to select a global path from any current position to the destination for an ASV. By exploiting environmental data such as islands, terrains and buoys, provided in the map, an optimal and safe route in terms of short-distance and collision-free from static obstacles is searched. Route planning has been a long-standing research topic in the research communities of computer science and transportation science and hence is quite well-studied. Typical strategies include discretisation of the map first followed by the application of motion planning algorithms to solve the coined graph search problem [22,23]. A number of classic motion planning algorithms including Dijkstra's, Rapidly-exploring Random Trees (RRT), and A* have been adapted and integrated into various assistive navigation systems of manned marine vehicles. Without much modification, these path planning methods can also be adopted in ASVs

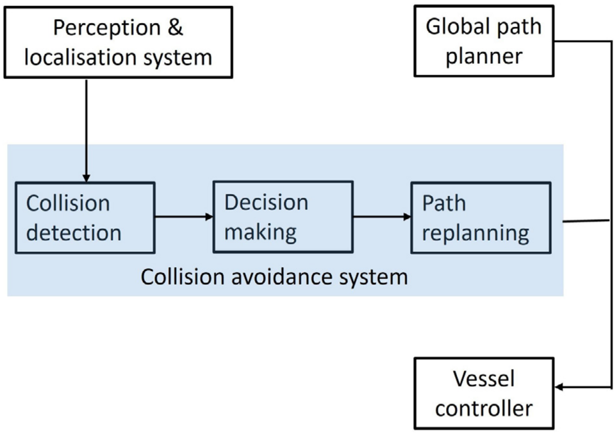


Fig. 2. The navigation, guidance and control system of ASVs.

easily. Since the marine route planning technique is quite mature, we will not discuss it in detail in the paper and readers who are interested in it are referred to the comprehensive survey papers [9,24]. A comparison of route planning and path replanning in the CAS will be covered in Section 3.4 where path replanning methods are reviewed.

2.1.3. Collision avoidance

The objective of CASs is to ensure ASVs' safety at all times despite encountering moving vessels or other obstacles in the vicinity. The CAS is generally composed of three functional modules of risk detection, decision making and path replanning. The CAS receives localisation and motion information such as the OS and TSs' position, speed and heading from the perception and localisation system continuously. The risk detection module regularly evaluates collision risks with the surrounding vessels and/or landmass and, if confirmed, will trigger an alarm for subsequent decision making and path replanning modules. Once a collision risk is deemed to exist, the decision making module will determine whether the TS's behaviour complies with COLREGs rules and which COLREGs encounter, i.e., "head-on", "crossing" or "overtaking", should be applied (The COLREGs rules will be summarised in the next subsection). Based on the decision made, the ASV re-plans its path locally over the existing global route such that the risk of collision is eliminated. In the next section, each submodule of the CAS will be discussed in sequence.

2.1.4. Vessel control

The path planning and collision avoidance decisions are executed by the ASV through vessel control systems. Typically, a sequence of waypoints or a reference trajectory is generated by the global and local path planners, and passed to vessel controller. With different control objectives, ASV control tasks can be classified into setpoint tracking, path following and trajectory tracking. Many nonlinear control methods have been employed for ASV control. Due to under-actuated dynamics of ASV, environment disturbance such as wave and wind, and actuator saturation, control of ASVs remains a challenging and active research topic. Since our attention in this review is focused on collision avoidance and path replanning strategies, vehicle control is not covered. Any reader interested in ASV control methods are referred to the comprehensive survey in [2,25].

2.2. Regulations at sea

The COLREGs were set out in 1972 by the IMO as a set of guidelines for vessel encounters at sea. It is expected that all vessel operators comply with these regulations, which outline procedures for determining right of way and correct avoidance manoeuvres. Without a human operator or crew physically present onboard the vessel, an autonomous ship must still obey the rules if it is to be lawfully operational at sea. Otherwise unpredictable or incorrect actions may lead to confusion and potentially catastrophic collisions amongst other marine traffic. The regulations are comprised of three main sections:

- General rules (Part A): outline the applicability and responsibilities of the regulations.
- Steering and sailing rules (Part B): consist of two sections, where Section 1 refers to the conduct of vessels in any visibility conditions and Section 2 regards the conduct of vessels in sight of one another.
- Lights and shapes (Part C): cover protocols for issuing indicating, warning or distress signals etc. and safe guidelines for the use of lighting.

Most early research focused on simple 1-1 ship encounter scenarios dictated by rules 13–16 in Part B. Later on, researchers have considered more complex scenarios, such as multi-ship encounter, COLREGs non-compliant TSs, operation in traffic-heavy area dictated by the traffic separation scheme (TSS), and in restricted visibility. For these scenarios, rules 2(b), 8, 10, 17 and 19 rather than rule 13–16 should be considered. All the relevant rules are rephrased briefly below. For easy referring the rules to the textbook such as the textbook [10], the original COLREGs rule numbers are kept:

- Rule 2(b) - Responsibility: under special circumstances, a departure from the rules may be made to avoid immediate danger.
- Rule 8 - Actions to avoid collision: actions shall be made in ample time. If there is sufficient sea-room, alteration of course alone may be most effective. Reduce speed, stop or reverse if necessary. Action by a ship is required if there is a risk of collision, and when the ship has right-of-way.
- Rule 10 - TSS scheme. This rule consists of quite a number of items and so we omit its details due to space constraint and the reader is referred to the textbook [10] for more information.
- Rule 13 - Overtaking: any vessel overtaking any other shall keep out of the way of the vessel being overtaken.
- Rule 14 - Head-on Situation: When two power-driven vessels are meeting on reciprocal or nearly reciprocal courses so as to involve risk of collision, each shall alter her course to starboard so that each shall pass on the port side of the other.
- Rule 15 - Crossing Situation: when two power-driven vessels are crossing so as to involve risk of collision, the vessel which has the other on her own starboard side shall keep out of the way and shall, if the circumstances of the case admit, avoid crossing ahead of the other vessel.
- Rule 16 - Actions by give-way vessel: every vessel which is directed to keep clear of another vessel shall, so far as possible, take early and substantial action to keep well clear.
- Rule 17 - Actions by stand-on vessel: where one of two vessels is to keep out of the way, the other shall keep her course and speed. The latter vessel may however take action to avoid collision by her manoeuvre alone, as soon as it becomes apparent to her that the vessel required to keep out of the way is not taking appropriate action in compliance with these rules.
- Rule 19 - Conduct of vessels in restricted visibility: vessel which detects by radar alone the presence of another vessel shall determine if a close quarters situation is developing and/or risk of collision exists. If so, she shall take avoiding action in ample time.

3. Collision avoidance and motion planning

Various collision avoidance and motion planning methods have been developed and applied in ASV navigation systems. In this section, we will first introduce three ASV models and comment how they are used for developing ASV navigation systems. Then we classify and review methods developed in the three stages of ASV navigation system one by one, namely collision detection, decision making and path replanning.

3.1. ASV modelling for motion prediction and planning

ASV motion model is needed when predicting the TS's motion and for planning the OS's path. Since ASVs typically travel in 2D space, the configuration space of an ASV is $\mathcal{R}^2 \times \mathcal{C}$, i.e, its position and orientation. Based on how the constraints and dynamics considered in modelling, several motion models have been proposed:

3.1.1. Constant-velocity model [26]

This is the simplest model based on the assumption that the ASV moves with constant velocity. The model is described as follows:

$$\dot{x} = v_x, \quad \dot{y} = v_y \tag{1}$$

where x and y are the coordinates of the ASV, and $\dot{x}$ and $\dot{y}$ are the speed of the ASV in the x and y axis of the coordinate frame.

3.1.2. Planar kinematic model [27]

Adding an additional assumption of constant yaw rate, the constant velocity model is extended to the planar kinematic model, which is written as follows:

$$\dot{\eta} = R(\psi)\nu \tag{2}$$

where $\eta = [x, y, \psi]$ represents position and heading in the earth-fixed frame, $\nu = [v_x, v_y, r]$ denotes surge and sway relative velocities and yaw rate decomposed in the body-fixed frame, and $R(\psi)$ is the rotation matrix from body-fixed to earth-fixed frame.

3.1.3. 3-Dof dynamic model [27]

Unlike the previous two kinematic models, the 3-dof ship model considers the mass and dynamics of vessels. In the model, motion in the horizontal plane is characterised and roll, pitch and heave motion are all neglected, which is written as follows:

$$\begin{cases} \dot{\eta} = R(\psi)\nu + \nu_c \\ M\dot{\nu} + C(\nu)\nu + D(\nu)\nu = \tau + R(\psi)^T \tau_\omega \end{cases} \tag{3}$$

where $\eta = [x, y, \psi]$ represents position and heading in the earth-fixed frame, $\nu = [v_x, v_y, r]$ denotes surge and sway relative velocities and yaw rate decomposed in the body-fixed frame, M is the vessel inertia matrix, $C(\cdot)$ and $D(\cdot)$ model, respectively, Coriolis and damping terms, $R(\psi)$ is the rotation matrix from body-fixed to earth-fixed frame, the input τ represents the commanded thrust and moments, and ν_c is the ocean current velocity and τ_ω is the wind force, both expressed in the earth-fixed frame.

3.1.4. Discussion on model choice

Among the three models introduced above, the constant velocity model is widely used in predicting the TS's trajectory where prior knowledge of the TS such as the inertial dynamics is usually unavailable. That is, the TS is assumed to keep its speed and heading in the near future. The planar kinematic model is used for predicting the TS's trajectory in some scenarios where the TS is believed to change its heading [28,29]. For example, during the head-on scenario, both vessels involved are expected to move to their own starboard sides according to the COLREGs Rule 14. The 3-dof model is typically used in ASV controller design where precise dynamics information is needed [30,31].

3.2. Collision detection

To avoid the risk of immediate collision and late detection of potential collision, the ASV needs to keep monitoring of TSs in the surroundings and assess the collision risk continuously. The entire process of collision detection consists of two steps: motion prediction, and risk assessment. As depicted in Fig. 4, we classify motion prediction into model-based and learning-based methods, and risk assessment into CPA-based and ship domain based methods.

3.2.1. Motion prediction

The model based methods predict TSs' motion using ship models introduced in Section 3.1, while the learning based methods do not use explicit ship models but predict TSs' motion from the historical traffic data. Depending on the scenarios investigated, both the simple constant-velocity and the planar kinematic models have been adopted in model-based motion prediction. In [7,15,19], the constant velocity model is used for predicting the TS's trajectory in the near future. The planar kinematic model was used where the TS broadcasts its intention of manoeuvre to the OS [32,33]. The motion of vessels is inevitably affected by external disturbance such as the wind and current, and the measurements from sensors are subject to environmental noises. To address the uncertainty caused by model and measurement noises, state estimation algorithms such as Kalman filtering were also incorporated in model-based motion prediction [28,29].

However, if the intention and manoeuvre behaviour of TSs during close encounter is not known, the model-based methods may fail to estimate the TS's trajectory. This fact constitutes the major motivation of developing learning-based motion prediction methods. The key to predict TSs' motion accurately without a model is the availability of huge amount of real-world marine traffic data and advance machine learning techniques. From 2002, AIS has been made compulsory by the IMO for ships over 300 gross tonnages on international voyages, cargoes over 500 gross tonnages in all water and all passenger vessels. From statistics, about 250,000 vessels worldwide had been fitted with AIS by 2012 [34]. The AIS broadcasts AIS data, including ship location, speed, course, heading, rate of turn, destination and estimated arrival time, ship name, ship type, ship size, current time through the very high frequencies transceiver. So far, quite a few open-access AIS data sources have been provided online for research, as surveyed in [20]. Machine learning techniques use the historic AIS data to train a model, and then the trained model is used to predict TS's manoeuvre and trajectory in real-world ship encounter situations. Different machine learning algorithms such as neural networks [35], Gaussian process [36,37], dynamic Bayesian network [38] and hidden Markov model [39] were proposed for predicting the TS's motion.

3.2.2. Risk assessment

The closest point of approach (CPA) method and its variations have been widely adopted for evaluating if there is a potential collision risk in the near future for the ASV with a TS [14,15,19,40–42]. In the CPA method, based on the assumption that both the OS and TS maintain their velocities, two quantities called the DCPA and the TCPA are calculated, which is illustrated in Fig. 5. Given that an ASV at position A encounters a TS at position C, the DCPA (distance to CPA) and TCPA (time to CPA) are formulated as follows:

$$\text{DCPA} = \delta_{LOS} \sin\alpha$$
$$\text{TCPA} = \frac{\delta_{LOS}\cos\alpha}{V_{ref}}$$

where V_{ref} is the projected relative velocity vector, δ_{LOS} is the current distance between the ASV and target, and α is the angle from the vector V_{ref} to the relative orientation from the ASV to the TS.

The CPA-based risk assessment methods compare the TCPA and/or the DCPA with prescribed thresholds $t_{\min}$ and/or $d_{\min}$, where $t_{\min}$ and $d_{\min}$ are determined manually based on the vessel types and the environment where they are being operated.

A number of CPA-based risk criteria have been proposed. The simplest one just compared the value of DCPA with a prescribed safe distance $d_{\min}$. A risk of collision was deemed to exist if the former is smaller than the latter [14,29,40]. This criteria becomes unrealistic in traffic-dense areas where the vessel usually have a very small DCPA with other vessels travelling reciprocally even from quite long distances away but the vessel only manoeuvrers to avoid collision with TSs in its close range. Motivated by that, [19] proposed that a risk criteria should be composed of both the DCPA and TCPA. In order to consider the effect

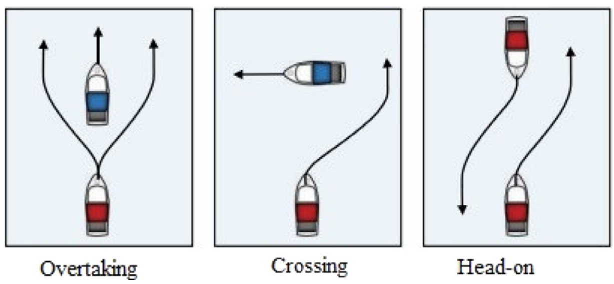


Fig. 3. Evasive behaviour defined by COLREGs rules in three typical encounter scenarios, where Red ships take an evasive maneuver (give-way) and blue ships maintain their course and speed (stand-on).
Source: Adapted from [38].

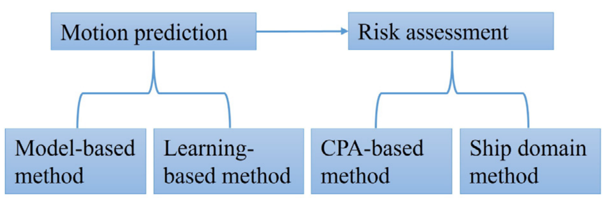


Fig. 4. The classification of collision detection in ASVs.

of uncertainty of motion prediction on collision risk, collision in the form of probability instead of in binary form were also presented in [29,43,44]. Monte Carlo method was employed in [29,43] to calculate the collision risk approximately, and a cross entropy method was proposed in [44] to estimate the probability of collision with low variance.

Beside the CPA, another risk assessment measure called the ship domain was proposed and have been used by navigators for collision risk assessment in practice. Compared with the CPA-based method where a circular safe zone around the OS is considered, the ship domain sets different safe distances in all directions around the OS. As a result, the safe region around the OS defined by ship domain is in an irregular geometric shape rather than a circle. As traditionally used by navigators sitting on manned ships, the ship domain is a subjective concept relying on experts' belief about the safe region. Therefore, how to extend the ship domain based risk assessment method for ASV navigation is still an open problem. A potential tool to construct the ship domain from historical traffic data is machine learning. [45] used neural networks to learn a representation of ship domain from simulation data. [46] proposed the concept of probabilistic ship domain and learning it from the widely available AIS data using the kernel density estimation method. [47] proposed a rough set based method to extract parameters that defined the shapes of ship domain.

3.3. Decision making

Once a collision risk is deemed to exist, the decision maker needs to determine (1) the specific COLREGs rule that should apply to the encounter, i.e., head-on, crossing or overtaking as shown in Fig. 3; (2). if any TS's behaviour breaks the COLREGs rules. The scene-division decision chart drafted by mariners has been widely used in practice for choosing COLREGs rules [42,48]. As shown in Fig. 6, the chart divides the plane into different COLREGs zones based on the relative position and bearing of the OS and TS. Given a ship encounter scenario and referring to the chart, it is uniquely determined which sector the TS falls in and which COLREGs rule applies accordingly.

If the uncertainty in sensor measurements and motion prediction is not considered properly, it might mislead to the wrong COLREGs rule selection. To overcome such a problem, several artificial intelligence techniques including fuzzy logic, Bayesian inference and neural network have been adopted together with the scene-division decision

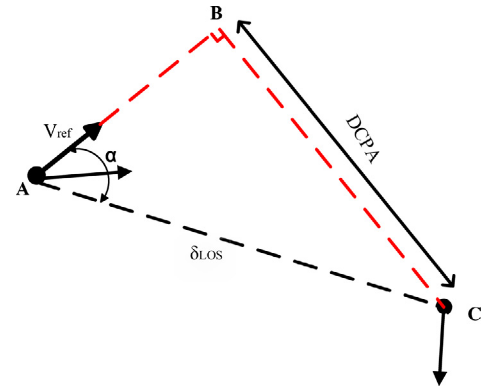


Fig. 5. The closest point of approach (CPA) for risk assessment.
Source: Adapted from [19].

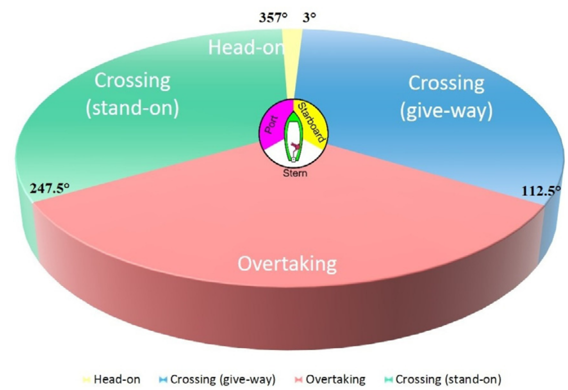


Fig. 6. The scene-division decision chart for COLREGs rules selection.

chart. Fuzzy logic was used in ASV decision making due to its ability to describe traffic rules linguistically and to deal with uncertainty [49–52]. The graphic model Bayesian network was introduced in [53] to make decision under motion uncertainty and in multi-vessel encounter situations. Neural networks was proposed to learn correct decision making from historical marine traffic data in [54], and semi-supervised deep neural networks was proposed in [55] to learn a classifier for ship encounter situation classification.

The scene-division decision chart method selects COLREGs rules for the OS based on the assumption that the TS follows COLREGs rules correctly. Nonetheless, some TSs' behaviours may violate the COLREGs rules sometimes. For example, consider that a TS supposed to be overtaking an ASV maintains her course and speed and closes dangerously upon the stern of the ASV. In this case, if the ASV maintains her course and speed as required by COLREGs rule 13 for "being overtaken" (please refer to Section 2.2 for COLREGs rules), a collision would occur soon. Actually, the decision maker should supersede the COLREGs rule 13 selected according to the decision chart and instead take an evasion action immediately to avoid potential collision. Note that both the CPA and ship domain methods fail to detect TS's COLREGs non-compliant behaviours as they cannot distinguish a collision risk due to a TS following COLREGs rules from the one that is not. To detect COLREGs non-compliant behaviour, [19] proposed an extended risk assessment criterion with two levels of risks where the urgent risk meant that the TS's behaviour violated COLREGs rules and created urgent risk of collision. [38] detected the TS's abnormal behaviours through estimating its manoeuvring intent and state evolution constantly using a probabilistic graphical model.

Table 1
Classification of path replanning methods.

Rule/behaviour based methods	[5,7,41]
Hybrid methods	A*-variants [17,40,56]; RRT-variants [57,58]
Reactive methods	Potential field method [59,60]; VO and its variants [15,48,61–65]
Optimisation-based method	Evolutionary algorithms [19,42,66–70]; MPC [30]
DRL approaches	[31,71–78]

3.4. Path replanning

Once a collision risk is detected and a give-way decision is made, an evasive trajectory should be planned immediately. This process is achieved by path replanning. Before review different path replanning techniques, we first explain the difference between path replanning and (route) planning as below:

(1) route planning considers only static obstacles, but path replanning considers moving vessels or obstacles;
(2) route planning usually is done offline, namely planning before a navigation task starts, but path replanning must be done online, that is planning an evasive trajectory on-the-fly once a give-way decision is made to remove collision risks;
(3) route planning generates a global path from the starting location to the destination while path replanning only modifies the route locally and return to the global path generated by route planning once collision risk is removed;
(4) route planning need not consider COLREGs rules, but path replanning need and it must ensure that the OS's trajectory is both safe and in accordance with COLREGs rules.

A variety of path replanning techniques with consideration of COLREGs rules have been developed for ASVs in recent years, which are categorised into five classes as shown in Table 1 in this review: rule and behaviour based methods; hybrid methods; reactive methods; optimisation-based methods and DRL-based methods. The former four methods are discussed in details below and the DRL base methods are left to the next section.

3.4.1. Rule/behaviour based methods

Early research on ASV collision avoidance primarily adopted rule or behaviour based methods to integrate COLREGs rules into path replanning. The seminar work [5,7] on COLREGs compliance proposed a behaviour-based local path planner. [41] proposed a manual biasing scheme that decided waypoints for the line-of-sight (LOS) guidance law to follow the rules of the road. Most rule-based methods involved hand-crafted design and focused on simple ship-encounter scenarios, and hence were hard to be extended for more complex ship encounter scenarios.

3.4.2. Hybrid methods

Hybrid methods refer to a class of path replanning methods that are improved over traditional route planning techniques. The methods improved upon classic route planning in two aspects. Firstly, it was faster than the original path planning algorithms such that it can be used for real-time CASs. Secondly, the specification of COLREGs-compliance was incorporated into path replanning algorithms seamlessly. Some hybrid methods modified traditional path planning algorithms such as A* and RRT. [40] proposed a R-RA* algorithm that improved over classic A* algorithm in that it only repaired the local, proximal path each time, thus saving computation time. [57] proposed COLREG-RRT, an RRT-based planner that was capable of identifying long-term, COLREGS-compliant trajectories with a high navigation success rate. Some other research added a suitable local path planner on the existing global path planner. The dynamic window algorithm (DWA) proposed for robot collision avoidance in [79] was used for local path planning, and combined with an A*-based global path planner in [58] and the theta* based global path planner in [56], respectively. In [17], the fuzzy inference algorithm was developed as local planner and then combined with an A*-based global planner.

3.4.3. Reactive methods

Reactive methods refer to a class of fast and reactive path planning algorithms that avoid collision by constructing repulsive fields or velocity obstacles around TSs. It includes artificial potential field (APF) methods, velocity obstacle (VO) methods, etc. APF generates a repulsive potential field around obstacles and an attractive potential field around the destination. The force from the potential field guides the motion of the vessels. The basic APF method can only deal with static obstacles [80]. To adapt APF for moving TSs and COLREGs-compliance, [59] introduced parameter adaptation and gain tuning schemes to the basic APF algorithm. Similarly, [60] added a virtual forces to APF to count for COLREGs rules for the OS.

VO was developed by researchers to solve real-time path planning for robotics in [81,82]. As to VO-based ASV navigation, COLREGs-compliance is enforced by integrating the forbidden zone into VO directly. The first-ever VO-based COLREGs-compliant ASV navigation system was tested successfully by on-water experiments [15]. Since then, VO methods have attracted the attention of many researchers. A number of VO-variant methods have been proposed and adapted for COLREGs-compliant navigation [48,61–65]. To relax the assumption of linear constant-speed motion of both the TS and OS, the non-linear velocity obstacle (NLVO) method was adopted in [62] for ASV path replanning, and adaptive RO method that took into consideration the kinematic and dynamic constraints of ASVs was proposed in [61]. To take account of uncertainty in motion prediction, the probabilistic velocity obstacle (PVO) method was introduced for ASV path replanning in [64]. To consider the possible interaction between TSs and the OS, optimal reciprocal velocity obstacle (RVO) method was introduced in [83] for cooperative ASV navigation. Furthermore, proactive strategies was incorporated into RVO such that cooperative behaviours between TSs and OSs were encouraged in [63]. [65] proposed a path planner that combined VO and RRT algorithms to generate safe trajectories for ASVs while following COLREGs rules.

3.4.4. Optimisation-based methods

ASV path replanning can be formulated as an optimisation problem as follows:

$$\min_{u} \quad J(x_{OS}, x_{TS}) \tag{4}$$

$$\text{s.t.:} \; g_s \leq 0 \tag{5}$$

$$g_r \leq 0 \tag{6}$$

$$x_{OS}(k+1) = f(x_{OS}(k), u(k), k) \tag{7}$$

$$x(k) \in \mathcal{X}, x(0) = x_0 \tag{8}$$

where the objective function to be optimised J is a function of the states of both the OS (x_{OS}) and TSs (x_{TS}). It can be a scalar or a vector function depending on the tasks and scenarios. For example, the object to be optimised can be the length of path, or both the length of path and the time to travel. In the former situation, it is a single-objective optimisation problem while in the latter situation, it becomes a multi-objective optimisation one. The collision avoidance requirement and COLREGEs rules are naturally formulated as constraints in terms of mathematical inequalities in (5) and (6), respectively. The dynamic constraint in (7) reflects the dynamics of ASV and it could be an optional choice. If the dynamic constraint (7) is considered then the optimisation problem becomes a dynamic optimisation one; otherwise it is a static optimisation problem. The last constraint (8) describes the kinodynamic constraints of the ASV such as speed limit and maximum turning rate.

Since the objective function for ASV path replanning is usually nonlinear and nonconvex, it is usually intractable to solve the optimisation problem analytically. The heuristic method called evolutionary algorithms have been borrowed to address ASV path replanning since they can solve complex static optimisation problems quite fast. With consideration of only a single optimisation objective, colony optimisation algorithm was adopted in [67], evolutionary programming in [69], and differential evolution in [70] to generate near-optimal COLREGs-compliant paths for ASVs. When multiple optimisation objectives were considered, [68] combined PID control law with a benchmark multi-objective optimisation algorithm NSGA II to control the ASV, and [19,42] proposed a hierarchical optimisation strategies incorporated into multi-objective particle swarm optimisation algorithm to prioritise safety over other objective functions. Similarly the lexicographic optimisation algorithm was adopted in [66] to optimise competing objectives hierarchically.

The model predictive control (MPC) method has been widely used to solve optimal planning and control problems for dynamic systems. Not until very recently the MPC has been introduced for ASV path replanning probably due to the relative high computation burden in MPC methods. To reduce computation burden, [30] tried to discretise the control input (course and speed changes) to a limited candidate values, which sped up the MPC for application in ASV real-time collision avoidance. Furthermore, the proposed MPC based path replanning method was tested successfully through on-water experiments [18,84]. In addition, distributed nonlinear MPC was introduced to handle multi-vessel cooperative path replanning in [85,86].

4. Deep reinforcement learning approach for safe navigation

The field of deep reinforcement learning (DRL) has seen rapid development over the last few years, as evidenced by many impressive achievements such as playing chess, playing video games and robot control at a superhuman level [87–89]. Motivated by that, researchers have introduced DRL techniques for ASV collision avoidance and path replanning. Different from traditional path replanning methods, DRL methods are generally model-free and they learn an end-to-end policy from observations to actions through the principle of trial and error.

4.1. Deep reinforcement learning

This section gives a brief overview of the concepts of DRL required in the review. The purpose is not to give a complete detail of the concepts and hence the descriptions are short. However, references to the resources are provided when deemed necessary.

In reinforcement learning, the task is generally formulated as a Markov decision process (MDP), defined as a tuple $\langle \mathcal{S}, \mathcal{A}, \mathcal{R}, \mathcal{P}, \gamma \rangle$ [90]. Here $\mathcal{S}$ is the state space, $\mathcal{A}$ is the action space, $\mathcal{R}$ is the reward space, $\gamma \in [0, 1)$ is the discount factor, $\mathcal{P}(s'|s,a)$ defines the transition probability distribution of the next new state s' when taking action a at state s where $s, s' \in \mathcal{S}, a \in \mathcal{A}$, and $R_{t+1}(s_t, a_t)$ denotes the instantaneous reward obtained by taking action a_t at the state s_t. Just to note, Cost instead of reward is often used in some control literature where maximisation in reinforcement learning setup is changed to minimisation instead.

Considering ASV path replanning as a MDP, the ASV takes an action a_t (speed and heading changes, acceleration commands) at the state s_t (location, speed and heading at time instant t), transits to the next state s_{t+1} following the transition dynamics $\mathcal{P}(s_{t+1}|s_t, a_t)$, and receives a reward $R_{t+1}(s_t, a_t)$. The aim of reinforcement learning algorithms is to learn a good policy that maximises the expected discount return in ASV path replanning. The discount return is defined as follows:

$$G_t = R_t + \gamma R_{t+1} + \cdots + \gamma^N R_{t+N}$$

where N is the number of steps in an episode. If a stochastic policy $\pi(a|s)$ is learnt, then the action will be sampled from the stochastic policy, that is $a \sim \pi(a|s)$ for a given state s; if a deterministic policy $\mu(s)$ is learnt, then actions will be selected deterministically for a given state s. Two important concept called value function $V^\pi(s)$ and state–action value function $Q^\pi(s,a)$ in reinforcement learning are introduced as below:

$$V^\pi(s) = \mathbb{E}_\pi[G_t | s_t = s]$$
$$Q^\pi(s,a) = \mathbb{E}_\pi[G_t | s_t = s, a_t = a]$$

where the difference between $V^\pi(s)$ and $Q^\pi(s,a)$ lies in that the former one considers the expected return when starting with state s and following the policy π all the time, and the latter one considers the expected return when starting with taking action a at state s and then following the policy π thereafter.

Furthermore, the optimal state value function and state–action value function and the optimal policy are defined as follows:

$$V^*(s) = \max_\pi V^\pi(s)$$
$$Q^*(s,a) = \max_\pi Q^\pi(s,a)$$
$$\pi^*(a|s) = \operatorname*{argmax}_a Q^*(s,a)$$

The great successes of deep neural networks in machine learning have advanced a new reinforcement learning paradigm DRL, where deep neural networks are employed as powerful nonlinear approximators of the optimal value and policy functions. Recently, the research on DRL have soared and a number of DRL algorithms have been invented. It is impossible to review all DRL algorithms in this subsection, so we will only briefly list the most influential DRL algorithms that are widely adapted for ASV path replanning. So far, all the developed DRL algorithms can be classified into three classes:

- Value-based methods: This kind methods estimate the optimal values of all different states, then derive the optimal policy using the estimated values. Widely used value-based methods include DQN [87] and other DQN-variants such as Double DQN [91], Dueling DQN [92], Rainbow [93].
- Policy-based methods: This kind of methods optimise the policy directly without maintaining the value functions. The most popular policy-based methods include A3C/A2C [94], PPO [95] and TRPO [96].
- Actor–critic methods: This kind of methods can be viewed as a combination of the above two methods. They maintain an explicit representation of both the policy (the actor) and the value estimates (the critic). Typical actor–critic methods include DDPG [97], TD3 [98] and SAC [99].

The implementation of all the above-mentioned algorithms are provided in the open-access library OpenAI Baselines [100] and an improved implementation based on OpenAI Baselines called Stable Lines [101], which can be easily adapted for DRL applications.

4.2. Deep reinforcement learning methods for ASV navigation

We summarise the developments of various DRL based ASV path replanning methods from three aspects: simulation and test environments, DRL based ASV navigation methods, and reward functions.

4.2.1. Simulation and test environments

Even though DRL is a model-free method, we still need an accurate model to generate simulation data for training. As to ASV navigation tasks, both the ASV dynamics and ocean environment make up of the environment. The training environment should be diverse enough such that the learnt policy could be generalised to other unseen complex environments [74]. Typically, the ASV simulator was built by the researchers using the 3-DOF ship model (3), as used in [31,72–74,77,78]. [102] published an open-access ASV simulation framework using the Python library and OpenAI Gym. Some other researchers use real-world data rather than simulation data for training. [103] used pre-recorded vessel trajectory dataset and [104] used the real AIS data as training data. As to test, most research still relied on computer simulation, except [72,105] used real ASVs deployed with the learnt DRL algorithms for experimental tests.

4.2.2. DRL based ASV navigation methods

So far, all three different kinds of DRL algorithms mentioned in Section 4.1 have been adapted for ASV path replanning. The standard DQN algorithm was adapted in [72,106], and two value-based DRL algorithms DQN and double DQN were tailored in [71,105] for ASV path planning and collision avoidance. [73] used the PPO algorithm to learn multi-ship collision avoidance method. Similarly, [74] applied PPO algorithm to learn avoid collision policy with both static and dynamic obstacles when following a given global path. Furthermore, [75] used the PPO algorithm from [74] and hand-crafted a reward function that encourages the ASV to comply with the COLREGs rules. [77] combined long short term memory (LSTM) with PPO algorithms to better retrieve temporal information such as waypoint sequences in ASV navigation. [78,104] proposed different modified DDPG algorithms to learn path replanning policy for collision avoidance. [31] added a soft actor–critic (SAC) algorithm that account for both model uncertainty and collision avoidance on a model-based controller. [103] proposed an inverse reinforcement learning algorithm to learn intelligent navigation behaviours from demonstration of real-world trajectories of manually-operated ships.

4.2.3. Reward functions

Since reinforcement learning algorithms learn an optimal policy that maximises the expected accumulated reward, reward function plays an important role in algorithm design of reinforcement learning. An appropriate reward function should exploit the domain knowledge of problems of concern. As to collision avoidance and path replanning, a basic idea is to construct a reward function as a combination of two parts: the reward of path following and that of collision avoidance [71–75]. The former one encourages the ASV to follow the planned path and the latter one penalises collision with other TSs. However, it is quite challenging to fine tune the weighting factor between the two reward terms and, in practice, it is usually obtained by trial and error [71]. One common issue with the reward function is sparsity, that is a reward can be obtained only at each episode's termination such as at the moment an ASV colliding with other vessels. Sparsity induces a very limited amount of feedback signals from the simulation and usually results in slow or non-convergent training in environments with long time horizons or challenging exploration. To address it, one needs to design the reward function manually or using reward shaping techniques. [107] compared different DRL algorithms for ASV navigation tasks and found that the reward function had important impact on learning performance and using reward shaping DRL algorithms with better performance were obtained. In addition, [103] circumvented manual design of reward function by learning it using an inverse reinforcement learning approach where the reward function was learnt from a dataset of prerecorded vessel trajectories using a kernel density estimation scheme.

4.2.4. DRL vs traditional path replanning methods

We compare DRL based and traditional path replanning methods and list their pros and cons as follows:

- The DRL based methods learn an end-to-end policy that maps from the perception to action directly for ASV navigation, without a need of collision detection and decision making modules as in traditional ASV CASs;
- The DRL based methods embody high-level intelligence such that the ASV's behaviour is not only COLREG-compliant but also of good seamanship. Since there are quite a number of COLREGs rules in total, it takes great effort to accommodate all COLREGs rule in decision making and path replanning modules of traditional ASV CASs, not to mention good seamanship. However, the DRL based method learn COLREGs-compliant behaviours and good seamanship directly from training data of ship encounters in various environments;
- Compared with traditional path replanning methods, the DRL based navigation techniques still lack of explainability and transparency. As a result, it is hard to provide a safe guarantee for DRL based navigation (at least by now), even though its performance could be quite good during test experiments. In addition, compared with traditional path replanning methods, it is much harder to tune the hyper-parameters in DRL algorithms.

5. Conclusion and future works

In this paper, we have reviewed recent advances on COLREGs-compliant navigation for ASVs in complex and dynamic environments. The three interconnected modules of ASV CASs (i.e., collision detection, decision making and path replanning) have been reviewed one by one. Both traditional methods and learning based methods are covered. We have analysed the motivation for developing learning based method, discussed the advantages/disadvantages of each methods and provided overviews on the corresponding research results. Based on the literature review, some related topics for the future research work are provided as follows:

- *Public simulation benchmarks for test:* So far, numerous ASV path replanning algorithms have been proposed, however, performance comparison of different algorithms are still rare. Simulation scenarios for test were usually designed by researchers in an ad hoc manner. To conduct transparent and fair comparison among algorithms, a public simulation benchmark that consists of a variety of typical ship encounter scenarios is desirable. With such a benchmark, the effectiveness of each newly-proposed algorithm can be justified easily by researchers.
- *Provable safety guarantees:* The effectiveness of path replanning algorithms were usually validated by simulation or real-world experiments at sea. However, in either approach, only a limited number of ship encounter scenarios can be tested, which leave a gap for real world application where safe navigation in all possible ship encounter scenarios are needed. As a result, it is an important yet challenging research topic to provide a provable safety guarantees for an ASV replanning algorithm.
- *Integration of traditional and learning methods:* As summarised in the paper, both the traditional and learning methods has been investigated for ASV path replanning. As each method has its own merit and cannot be substituted by the other one, it is a promising research direction that integrates the two methods seamlessly in ASV CASs, such that the powers of both methods are best exploited.
- *ASV navigation on urban waterways:* So far, almost all research on ASV navigation has focused to ocean environment. However, a demand of ASVs used for transporting people and delivering goods through the inland waterways such as canal, river and lakes is emerging [108,109]. Different from the sea and ocean, the urban waterways provide a much narrow, busy environment where COLREGs rules do not apply. How to extend existing research on safe navigation of ASVs at sea to that in the urban waterway is still an rarely-investigated area, except some very recent results [66,103,110–114], which could be another promising research direction in the future.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] S. Campbell, W. Naeem, G.W. Irwin, A review on improving the autonomy of unmanned surface vehicles through intelligent collision avoidance manoeuvres, Annu. Rev. Control 36 (2) (2012) 267–283.

[2] Z. Liu, Y. Zhang, X. Yu, C. Yuan, Unmanned surface vehicles: An overview of developments and challenges, Annu. Rev. Control 41 (2016) 71–93.

[3] A. Vagale, R. Oucheikh, R.T. Bye, O.L. Osen, T.I. Fossen, Path planning and collision avoidance for autonomous surface vehicles I: A review, J. Mar. Sci. Technol. (2021) 1–15.

[4] H.-C. Burmeister, M. Constapel, Autonomous collision avoidance at sea: A survey, Front. Robot. AI (2021) 297.

[5] M.R. Benjamin, Multi-objective autonomous vehicle navigation in the presence of cooperative and adversarial moving contacts, in: OCEANS'02 MTS/IEEE, Vol. 3, IEEE, 2002, pp. 1878–1885.

[6] M. Caccia, Autonomous surface craft: Prototypes and basic research issues, in: 14th Mediterranean Conference on Control and Automation, 2006, pp. 1–6.

[7] M.R. Benjamin, J.J. Leonard, J.A. Curcio, P.M. Newman, A method for protocol-based collision avoidance between autonomous marine surface craft, J. Field Robotics 23 (5) (2006) 333–346.

[8] T. Statheros, G. Howells, K.M. Maier, Autonomous ship collision avoidance navigation concepts, technologies and techniques, J. Navig. 61 (01) (2008) 129–142.

[9] C. Tam, R. Bucknall, A. Greig, Review of collision avoidance and path planning methods for ships in close range encounters, J. Navig. 62 (03) (2009) 455–476.

[10] A.N. Cockcroft, J.N.F. Lameijer, Guide to the Collision Avoidance Rules, Butterworth-Heinemann, 2003.

[11] U. Commandant, International Regulations for Prevention of Collisions at Sea, 1972 (72 COLREGS), Vol. 16673, US Department of Transportation, US Coast Guard, COMMANDANT INSTRUCTION M, 1999.

[12] W. Naeem, T. Xu, R. Sutton, A. Tiano, The design of a navigation, guidance, and control system for an unmanned surface vehicle for environmental monitoring, Proc. Inst. Mech. Eng., Part M: J. Eng. Marit. Environ. 222 (2) (2008) 67–79.

[13] L.P. Perera, V. Ferrari, F.P. Santos, M.A. Hinostroza, C.G. Soares, Experimental evaluations on ship autonomous navigation and collision avoidance by intelligent guidance, IEEE J. Ocean. Eng. 40 (2) (2015) 374–387.

[14] I.R. Bertaska, B. Shah, K. von Ellenrieder, P. Švec, W. Klinger, A.J. Sinisterra, M. Dhanak, S.K. Gupta, Experimental evaluation of automatically-generated behaviors for USV operations, Ocean Eng. 106 (2015) 496–514.

[15] Y. Kuwata, M.T. Wolf, D. Zarzhitsky, T.L. Huntsberger, Safe maritime autonomous navigation with colregs, using velocity obstacles, IEEE J. Ocean. Eng. 39 (1) (2014) 110–119.

[16] R. Fiskin, O. Atik, H. Kisi, E. Nasibov, T.A. Johansen, Fuzzy domain and meta-heuristic algorithm-based collision avoidance control for ships: Experimental validation in virtual and real environment, Ocean Eng. 220 (2021) 108502.

[17] J. Han, Y. Cho, J. Kim, J. Kim, N.-s. Son, S.Y. Kim, Autonomous collision detection and avoidance for ARAGON USV: Development and field tests, J. Field Robotics 37 (6) (2020) 987–1002.

[18] D.K.M. Kufoalor, T.A. Johansen, E.F. Brekke, A. Hepsø, K. Trnka, Autonomous maritime collision avoidance: Field verification of autonomous surface vehicle behavior in challenging scenarios, J. Field Robotics 37 (3) (2020) 387–403.

[19] L. Hu, W. Naeem, E. Rajabally, G. Watson, T. Mills, Z. Bhuiyan, C. Raeburn, I. Salter, C. Pekcan, A multiobjective optimization approach for COLREGs-compliant path planning of autonomous surface vehicles verified on networked bridge simulators, IEEE Trans. Intell. Transp. Syst. 21 (3) (2019) 1167–1179.

[20] E. Tu, G. Zhang, L. Rachmawati, E. Rajabally, G.-B. Huang, Exploiting AIS data for intelligent maritime navigation: A comprehensive survey from data to methodology, IEEE Trans. Intell. Transp. Syst. 19 (5) (2017) 1559–1582.

[21] Y. Huang, L. Chen, P. Chen, R.R. Negenborn, P. Van Gelder, Ship collision avoidance methods: State-of-the-art, Saf. Sci. 121 (2020) 451–473.

[22] H. Niu, Y. Lu, A. Savvaris, A. Tsourdos, An energy-efficient path planning algorithm for unmanned surface vehicles, Ocean Eng. 161 (2018) 308–321.

[23] H. Niu, Z. Ji, A. Savvaris, A. Tsourdos, Energy efficient path planning for unmanned surface vehicle in spatially-temporally variant environment, Ocean Eng. 196 (2020) 106766.

[24] M. Christiansen, K. Fagerholt, D. Ronen, Ship routing and scheduling: Status and perspectives, Transp. Sci. 38 (1) (2004) 1–18.

[25] Z. Peng, J. Wang, D. Wang, Q.-L. Han, An overview of recent advances in coordinated control of multiple autonomous surface vehicles, IEEE Trans. Ind. Inform. 17 (2) (2020) 732–745.

[26] X.R. Li, V.P. Jilkov, Survey of maneuvering target tracking. Part I. Dynamic models, IEEE Trans. Aerosp. Electron. Syst. 39 (4) (2003) 1333–1364.

[27] T.I. Fossen, Handbook of Marine Craft Hydrodynamics and Motion Control, John Wiley & Sons, 2011.

[28] L.P. Perera, P. Oliveira, C.G. Soares, Maritime traffic monitoring based on vessel detection, tracking, state estimation, and trajectory prediction, IEEE Trans. Intell. Transp. Syst. 13 (3) (2012) 1188–1200.

[29] B.C. Shah, P. Švec, I.R. Bertaska, A.J. Sinisterra, W. Klinger, K. von Ellenrieder, M. Dhanak, S.K. Gupta, Resolution-adaptive risk-aware trajectory planning for surface vehicles operating in congested civilian traffic, Auton. Robots 40 (7) (2016) 1139–1163.

[30] T.A. Johansen, T. Perez, A. Cristofaro, Ship collision avoidance and COLREGS compliance using simulation-based control behavior selection with predictive hazard assessment, IEEE Trans. Intell. Transp. Syst. 17 (12) (2016) 3407–3422.

[31] Q. Zhang, W. Pan, V. Reppa, Model-reference reinforcement learning for collision-free tracking control of autonomous surface vehicles, IEEE Trans. Intell. Transp. Syst. (2021).

[32] L. Chen, Y. Huang, H. Zheng, H. Hopman, R. Negenborn, Cooperative multi-vessel systems in urban waterway networks, IEEE Trans. Intell. Transp. Syst. 21 (8) (2019) 3294–3307.

[33] R. Scheepens, H. van de Wetering, J.J. van Wijk, Contour based visualization of vessel movement predictions, Int. J. Geogr. Inf. Sci. 28 (5) (2014) 891–909.

[34] Automatic identification system, 2022, [Online]. Available: https://en.wikipedia.org/wiki/Automatic-identification-system.

[35] T. Zhang, X.-Q. Zheng, M.-X. Liu, Multiscale attention-based LSTM for ship motion prediction, Ocean Eng. 230 (2021) 109066.

[36] J. Meng, Y. Liu, R. Bucknall, W. Guo, Z. Ji, Anisotropic GPMP2: A fast continuous-time Gaussian processes based motion planner for unmanned surface vehicles in environments with ocean currents, IEEE Trans. Autom. Sci. Eng. (2022).

[37] H. Rong, A. Teixeira, C.G. Soares, Ship trajectory uncertainty prediction based on a Gaussian process model, Ocean Eng. 182 (2019) 499–511.

[38] Y. Cho, J. Kim, J. Kim, Intent inference-based ship collision avoidance in encounters with rule-violating vessels, IEEE Robot. Autom. Lett. 7 (1) (2021) 518–525.

[39] G. Pallotta, S. Horn, P. Braca, K. Bryan, Context-enhanced vessel prediction based on ornstein-uhlenbeck processes using historical AIS traffic patterns: Real-world experimental results, in: 17th International Conference on Information Fusion, FUSION, IEEE, 2014, pp. 1–7.

[40] S. Campbell, M. Abu-Tair, W. Naeem, An automatic COLREGs-compliant obstacle avoidance system for an unmanned surface vehicle, Proc., Inst. Mech. Eng., Part M: J. Eng. Marit. Environ. 228 (2) (2014) 108–121.

[41] W. Naeem, G.W. Irwin, A. Yang, COLREGs-based collision avoidance strategies for unmanned surface vehicles, Mechatronics 22 (6) (2012) 669–678.

[42] L. Hu, W. Naeem, E. Rajabally, G. Watson, T. Mills, Z. Bhuiyan, I. Salter, Colregs-compliant path planning for autonomous surface vehicles: A multiobjective optimization approach, IFAC-PapersOnLine 50 (1) (2017) 13662–13667.

[43] P. Agrawal, J.M. Dolan, COLREGS-compliant target following for an unmanned surface vehicle in dynamic environments, in: Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on, IEEE, 2015, pp. 1065–1070.

[44] T. Tengesdal, T.A. Johansen, E.F. Brekke, Ship collision avoidance utilizing the cross-entropy method for collision risk assessment, IEEE Trans. Intell. Transp. Syst. (2021).

[45] J.-H. Ahn, K.-P. Rhee, Y.-J. You, A study on the collision avoidance of a ship using neural networks and fuzzy logic, Appl. Ocean Res. 37 (2012) 162–173.

[46] L. Zhang, Q. Meng, Probabilistic ship domain with applications to ship collision risk assessment, Ocean Eng. 186 (2019) 106130.

[47] J. Zhou, C. Wang, A. Zhang, A COLREGs-based dynamic navigation safety domain for unmanned surface vehicles: A case study of Dolphin-I, J. Mar. Sci. Eng. 8 (4) (2020) 264.

[48] W. Shaobo, Z. Yingjun, L. Lianbo, A collision avoidance decision-making system for autonomous ship based on modified velocity obstacle method, Ocean Eng. 215 (2020) 107910.

[49] S.-L. Kao, K.-T. Lee, K.-Y. Chang, M.-D. Ko, A fuzzy logic method for collision avoidance in vessel traffic service, J. Navig. 60 (01) (2007) 17–31.

[50] B. Wu, T. Cheng, T.L. Yip, Y. Wang, Fuzzy logic based dynamic decision-making system for intelligent navigation strategy within inland traffic separation schemes, Ocean Eng. 197 (2020) 106909.

[51] L. Perera, J. Carvalho, C.G. Soares, Fuzzy logic based decision making system for collision avoidance of ocean navigation under critical collision conditions, J. Mar. Sci. Technol. 16 (1) (2011) 84–99.

[52] L. Perera, J. Carvalho, C. Guedes Soares, Solutions to the failures and limitations of mamdani fuzzy inference in ship navigation, IEEE Trans. Veh. Technol. 63 (4) (2014) 1539–1554.

[53] L.P. Perera, J. Carvalho, C.G. Soares, Intelligent ocean navigation and fuzzy-Bayesian decision/action formulation, IEEE J. Ocean. Eng. 37 (2) (2012) 204–219.

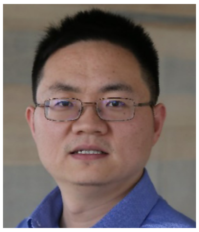
[54] T. Praczyk, Neural anti-collision system for autonomous surface vehicle, Neurocomputing 149 (2015) 559–572.

[55] X. Chen, Y. Liu, K. Achuthan, X. Zhang, J. Chen, A semi-supervised deep learning model for ship encounter situation classification, Ocean Eng. 239 (2021) 109824.

[56] N. Wang, H. Xu, Dynamics-constrained global-local hybrid path planning of an autonomous surface vehicle, IEEE Trans. Veh. Technol. 69 (7) (2020) 6928–6942.
[57] H.-T.L. Chiang, L. Tapia, COLREG-RRT: An RRT-based COLREGS-compliant motion planner for surface vehicle navigation, IEEE Robot. Autom. Lett. 3 (3) (2018) 2024–2031.
[58] P. Tang, R. Zhang, D. Liu, L. Huang, G. Liu, T. Deng, Local reactive obstacle avoidance approach for high-speed unmanned surface vehicle, Ocean Eng. 106 (2015) 128–140.
[59] W. Naeem, S. Campbell, L. Hu, A reactive COLREGs-compliant navigation strategy for autonomous maritime navigation, IFAC Proc. Vol. (2016).
[60] H. Lyu, Y. Yin, COLREGS-constrained real-time path planning for autonomous ships using modified artificial potential fields, J. Navig. 72 (3) (2019) 588–608.
[61] D.F. Campos, A. Matos, A.M. Pinto, An adaptive velocity obstacle avoidance algorithm for autonomous surface vehicles, in: 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, IEEE, 2019, pp. 8089–8096.
[62] Y. Huang, L. Chen, P. Van Gelder, Generalized velocity obstacle algorithm for preventing ship collisions at sea, Ocean Eng. 173 (2019) 142–156.
[63] D.K.M. Kufoalor, E.F. Brekke, T.A. Johansen, Proactive collision avoidance for ASVs using a dynamic reciprocal velocity obstacles method, in: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, IEEE, 2018, pp. 2402–2409.
[64] Y. Cho, J. Han, J. Kim, Efficient COLREG-compliant collision avoidance in multi-ship encounter situations, IEEE Trans. Intell. Transp. Syst. (2020).
[65] R. Dubey, S.J. Louis, VORRT-COLREGs: A hybrid velocity obstacles and RRT based COLREGs-compliant path planner for autonomous surface vessels, 2021, arXiv preprint arXiv:2109.00862.
[66] T. Shan, W. Wang, B. Englot, C. Ratti, D. Rus, A receding horizon multi-objective planner for autonomous surface vehicles in urban waterways, in: 2020 59th IEEE Conference on Decision and Control, CDC, IEEE, 2020, pp. 4085–4092.
[67] A. Lazarowska, Ship's trajectory planning for collision avoidance at sea based on ant colony optimisation, J. Navig. 68 (2) (2015) 291–307.
[68] X. Qingyang, Multi-objective based course-keeping controller optimization of unmanned surface vehicle, in: 2014 33rd Chinese Control Conference, CCC, IEEE, 2014, pp. 7483–7486.
[69] R. Szlapczynski, Evolutionary sets of safe ship trajectories: A new approach to collision avoidance, J. Navig. 64 (01) (2011) 169–181.
[70] Y.-X. Zhao, W. Li, S. Feng, W.Y. Ochieng, W. Schuster, An improved differential evolution algorithm for maritime collision avoidance route planning, Abstr. Appl. Anal. (2014).
[71] S.Y. Luis, D.G. Reina, S.L.T. Marín, A deep reinforcement learning approach for the patrolling problem of water resources through autonomous surface vehicles: The Ypacarai lake case, IEEE Access 8 (2020) 204076–204093.
[72] H. Shen, H. Hashimoto, A. Matsuda, Y. Taniguchi, D. Terada, C. Guo, Automatic collision avoidance of multiple ships based on deep Q-learning, Appl. Ocean Res. 86 (2019) 268–288.
[73] L. Zhao, M.-I. Roh, COLREGs-compliant multiship collision avoidance based on deep reinforcement learning, Ocean Eng. 191 (2019) 106436.
[74] E. Meyer, H. Robinson, A. Rasheed, O. San, Taming an autonomous surface vehicle for path following and collision avoidance using deep reinforcement learning, IEEE Access 8 (2020) 41466–41481.
[75] E. Meyer, A. Heiberg, A. Rasheed, O. San, COLREG-compliant collision avoidance for unmanned surface vehicle using deep reinforcement learning, IEEE Access 8 (2020) 165344–165364.
[76] S.Y. Luis, D.G. Reina, S.L.T. Marín, A multiagent deep reinforcement learning approach for path planning in autonomous surface vehicles: The Ypacaraí lake patrolling case, IEEE Access 9 (2021) 17084–17099.
[77] R. Sawada, K. Sato, T. Majima, Automatic ship collision avoidance using deep reinforcement learning with LSTM in continuous action spaces, J. Mar. Sci. Technol. 26 (2) (2021) 509–524.
[78] C. Zhou, Y. Wang, L. Wang, H. He, Obstacle avoidance strategy for an autonomous surface vessel based on modified deep deterministic policy gradient, Ocean Eng. 243 (2022) 110–166.
[79] D. Fox, W. Burgard, S. Thrun, The dynamic window approach to collision avoidance, IEEE Robot. Autom. Mag. 4 (1) (1997) 23–33.
[80] O. Khatib, Real-time obstacle avoidance for manipulators and mobile robots, in: Autonomous Robot Vehicles, Springer, 1986, pp. 396–404.
[81] P. Fiorini, Z. Shiller, Motion planning in dynamic environments using velocity obstacles, Int. J. Robot. Res. 17 (7) (1998) 760–772.
[82] J. Van den Berg, M. Lin, D. Manocha, Reciprocal velocity obstacles for real-time multi-agent navigation, in: 2008 IEEE International Conference on Robotics and Automation, Ieee, 2008, pp. 1928–1935.
[83] Y. Zhao, W. Li, P. Shi, A real-time collision avoidance learning system for unmanned surface vessels, Neurocomputing 182 (2016) 255–266.
[84] I.B. Hagen, D.K.M. Kufoalor, E.F. Brekke, T.A. Johansen, MPC-based collision avoidance strategy for existing marine vessel guidance systems, in: 2018 IEEE International Conference on Robotics and Automation, ICRA, IEEE, 2018, pp. 7618–7623.
[85] L. Ferranti, R.R. Negenborn, T. Keviczky, J. Alonso-Mora, Coordination of multiple vessels via distributed nonlinear model predictive control, in: 2018 European Control Conference, ECC, IEEE, 2018, pp. 2523–2528.
[86] L. Chen, H. Hopman, R.R. Negenborn, Distributed model predictive control for vessel train formations of cooperative multi-vessel systems, Transp. Res. C 92 (2018) 101–118.
[87] V. Mnih, K. Kavukcuoglu, D. Silver, A.A. Rusu, J. Veness, M.G. Bellemare, A. Graves, M. Riedmiller, A.K. Fidjeland, G. Ostrovski, et al., Human-level control through deep reinforcement learning, Nature 518 (7540) (2015) 529–533.
[88] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, et al., Mastering the game of go without human knowledge, Nature 550 (7676) (2017) 354–359.
[89] J. Hwangbo, J. Lee, A. Dosovitskiy, D. Bellicoso, V. Tsounis, V. Koltun, M. Hutter, Learning agile and dynamic motor skills for legged robots, Science Robotics 4 (26) (2019) eaau5872.
[90] R.S. Sutton, A.G. Barto, Reinforcement Learning: An Introduction, MIT Press, 2018.
[91] H. Van Hasselt, A. Guez, D. Silver, Deep reinforcement learning with double q-learning, in: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 30, (no. 1) 2016.
[92] Z. Wang, T. Schaul, M. Hessel, H. Hasselt, M. Lanctot, N. Freitas, Dueling network architectures for deep reinforcement learning, in: International Conference on Machine Learning, PMLR, 2016, pp. 1995–2003.
[93] M. Hessel, J. Modayil, H. Van Hasselt, T. Schaul, G. Ostrovski, W. Dabney, D. Horgan, B. Piot, M. Azar, D. Silver, Rainbow: Combining improvements in deep reinforcement learning, in: Thirty-Second AAAI Conference on Artificial Intelligence, 2018.
[94] V. Mnih, A.P. Badia, M. Mirza, A. Graves, T. Lillicrap, T. Harley, D. Silver, K. Kavukcuoglu, Asynchronous methods for deep reinforcement learning, in: International Conference on Machine Learning, PMLR, 2016, pp. 1928–1937.
[95] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, O. Klimov, Proximal policy optimization algorithms, 2017, arXiv preprint arXiv:1707.06347.
[96] J. Schulman, S. Levine, P. Abbeel, M. Jordan, P. Moritz, Trust region policy optimization, in: International Conference on Machine Learning, PMLR, 2015, pp. 1889–1897.
[97] T.P. Lillicrap, J.J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, D. Wierstra, Continuous control with deep reinforcement learning, 2015, arXiv preprint arXiv:1509.02971.
[98] S. Fujimoto, H. Hoof, D. Meger, Addressing function approximation error in actor-critic methods, in: International Conference on Machine Learning, PMLR, 2018, pp. 1587–1596.
[99] T. Haarnoja, A. Zhou, P. Abbeel, S. Levine, Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor, in: International Conference on Machine Learning, PMLR, 2018, pp. 1861–1870.
[100] P. Dhariwal, C. Hesse, O. Klimov, A. Nichol, M. Plappert, A. Radford, J. Schulman, S. Sidor, Y. Wu, P. Zhokhov, Openai baselines, 2017, https://github.com/openai/baselines,
[101] Stable baselines, 2022, [Online]. Available: https://github.com/DLR-RM/stable-baselines3.
[102] A Python simulation framework for collision avoidance for unmanned surface vehicle using deep reinforcement learning, 2021, [Online]. Available: https://github.com/ThomasNLarsen/gym-auv.
[103] S. Park, M. Cap, J. Alonso-Mora, C. Ratti, D. Rus, Social trajectory planning for urban autonomous surface vessels, IEEE Trans. Robot. 37 (2) (2020) 452–465.
[104] S. Guo, X. Zhang, Y. Zheng, Y. Du, An autonomous path planning model for unmanned ships based on deep reinforcement learning, Sensors 20 (2) (2020) 426.
[105] J. Woo, N. Kim, Collision avoidance for an unmanned surface vehicle using deep reinforcement learning, Ocean Eng. 199 (2020) 107001.
[106] X. Zhou, P. Wu, H. Zhang, W. Guo, Y. Liu, Learn to navigate: Cooperative path planning for unmanned surface vehicles using deep reinforcement learning, IEEE Access 7 (2019) 165262–165278.
[107] T.N. Larsen, H.Ø. Teigen, T. Laache, D. Varagnolo, A. Rasheed, Comparing deep reinforcement learning algorithms' ability to safely navigate challenging waters, Front. Robot. AI 8 (2021).
[108] Urban waterways in global cities, Forum Rep., Chicago Council Global Affairs, Chicago, IL, USA, 2017.
[109] Roboat, 2021, [Online]. Available: http://roboat.org.
[110] W. Wang, B. Gheneti, L.A. Mateos, F. Duarte, C. Ratti, D. Rus, Roboat: An autonomous surface vehicle for urban waterways, in: 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, IEEE, 2019, pp. 6340–6347.
[111] W. Wang, T. Shan, P. Leoni, D. Fernández-Gutiérrez, D. Meyers, C. Ratti, D. Rus, Roboat II: A novel autonomous surface vessel for urban environments, in: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, IEEE, 2020, pp. 1740–1747.
[112] B.H. Gheneti, Reconfigurable Autonomous Surface Vehicles: Perception and Trajectory Optimization Algorithms (Ph.D. thesis), Massachusetts Institute of Technology, 2019.

[113] M.S. Tannum, J.H. Ulvensøen, Urban mobility at sea and on waterways in Norway, J. Phys.: Conf. Ser. 1357 (1) (2019) 012018.

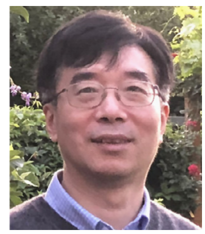
[114] Y. Cheng, M. Jiang, J. Zhu, Y. Liu, Are we ready for unmanned surface vehicles in inland waterways? The usvinland multisensor dataset and benchmark, IEEE Robot. Autom. Lett. 6 (2) (2021) 3964–3970.



Liang Hu received both the B.E. and M.E. degrees in Control Science and Engineering from Harbin Institute of Technology, China, in 2008 and 2010, respectively, and the Ph.D. degree in Computer Science from Brunel University London, U.K., in 2016. Dr Hu has been a lecturer in the School of Computer Science & Electronic Engineering at the University of Essex, United Kingdom from 2019 to 2022. Prior to that, He did the postdoctoral research at Queen's University Belfast and Loughborough University. His research interests include signal processing, control and decision and their applications in robotics and autonomous systems.

Huosheng Hu is a Professor in the School of Computer Science & Electronic Engineering at the University of Essex, United Kingdom. He received the M.Sc. degree in industrial automation from the Central South University in China and the Ph.D. degree in robotics from the University of Oxford in the U.K. His research interests include autonomous robotics, human–robot interaction, embedded systems, mechatronics, data fusion, learning algorithms, and industrial automation. He has published over 500 papers in journals, books, and conferences in these areas, and received several best paper awards. He has a citation of 17229 and his H-Index is 62 (Google Scholars). He has been recognised as among the top 100,000 or in the top 2% of most cited scholars in the world. Prof. Hu is a Fellow of Institute of Engineering & Technology, a Fellow of Institute of Measurement & Control, a life senior member of IEEE and a founding member of IEEE Robotics & Automation Society Technical committee on Networked Robots. He has been a Program Chair or a member of Advisory/Organising Committee for many international conferences such as IEEE IROS, ICRA, ICMA, ROBIO conferences. He is currently Executive Editor for International Journal of Mechatronics and Automation, as well as Former Editor-in-Chief for International Journal of Automation and Computing and Former Editor-in-Chief for MDPI Robotics Journal.

Wasif Naeem received the B.Eng. degree in Electrical Engineering from NED University of Engineering and Technology, Pakistan, in 1998, the M.Sc. degree in Electrical Engineering from King Fahd University, Saudi Arabia, in 2001, and the Ph.D. degree in Mechanical and Marine Engineering from the University of Plymouth in 2004. He joined Queen's University Belfast in 2007, where he is currently a Reader in robotics and autonomous systems. His expertise is in cooperative control, and indoor and outdoor robot navigation including mapping and collision avoidance, with applications to unmanned marine and aerial vehicles publishing over 100 peer-reviewed papers. He is a member of the IFAC Technical Committee on Marine Systems and the EPSRC College.



Zidong Wang was born in Jiangsu, China, in 1966. He received the B.Sc. degree in mathematics in 1986 from Suzhou University, Suzhou, China, and the M.Sc. degree in applied mathematics in 1990 and the Ph.D. degree in electrical engineering in 1994, both from Nanjing University of Science and Technology, Nanjing, China.

He is currently Professor of Dynamical Systems and Computing in the Department of Computer Science, Brunel University London, U.K. From 1990 to 2002, he held teaching and research appointments in universities in China, Germany and the UK. Prof. Wang's research interests include dynamical systems, signal processing, bioinformatics, control theory and applications. He has published more than 600 papers in international journals. He is a holder of the Alexander von Humboldt Research Fellowship of Germany, the JSPS Research Fellowship of Japan, William Mong Visiting Research Fellowship of Hong Kong.

Prof. Wang serves (or has served) as the Editor-in-Chief for International Journal of Systems Science, the Editor-in-Chief for Neurocomputing, the Editor-in-Chief for Systems Science & Control Engineering, and an Associate Editor for 12 international journals including IEEE Transactions on Automatic Control, IEEE Transactions on Control Systems Technology, IEEE Transactions on Neural Networks, IEEE Transactions on Signal Processing, and IEEE Transactions on Systems, Man, and Cybernetics-Part C. He is a Member of the Academia Europaea, a Member of the European Academy of Sciences and Arts, an Academician of the International Academy for Systems and Cybernetic Sciences, a Fellow of the IEEE, a Fellow of the Royal Statistical Society and a member of program committee for many international conferences.