King's Research Portal

DOI:

[10.1108/IR-08-2020-0160](https://doi.org/10.1108/IR-08-2020-0160)

Document Version

Peer reviewed version

[Link to publication record in King's Research Portal](#)

Citation for published version (APA):

Wen, S., Lv, X., Lam, H. K., Fan, S., Yuan, X., & Chen, M. (Accepted/In press). Probability Dueling DQN active visual SLAM for autonomous navigation in indoor environment. *Industrial Robot*, 48(3), 359-365.
<https://doi.org/10.1108/IR-08-2020-0160>

Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Probability Dueling DQN Active Visual SLAM for Autonomous Navigation in Complex Environment

Shuhuan Wen^a, Xiaohan Lv^a, Hak Keung Lam^{b,*}, Shaokang Fan^a, Xiao Yuan^a, Ming Chen^b

^a*Key Lab of Industrial Computer Control Engineering of Hebei Province, Yanshan University, Qinhuangdao, 066004, China*

^b*Department of Engineering, Strand, London, WC2R 2LS, United Kingdom*

Abstract

This paper presents an active SLAM framework for performing autonomous navigation under a complex environment with static and dynamic obstacles. We integrate a path planning algorithm with visual SLAM in order to decrease navigation uncertainty and build an environment map. We employ Monodepth method for improving the predictive speed of recognizing the obstacles and propose a Probability Dueling DQN (PD3QN) algorithm to optimize the path of the robot, which can reach the destination more quickly than Dueling DQN. Then PD3QN-based path planning algorithm is integrated with FastSLAM to achieve the autonomous navigation and map the environment. We show that the proposed method offers good performance over existing Dueling DQN for navigation uncertainty under the complex environment with different numbers of the static and dynamic obstacles in real world field. .

Keywords: Active SLAM, Deep Reinforcement Learning, Path Planning, FastSLAM

*The work is supported by the national Natural Science Foundation of China (Project No. 61773333).

*Corresponding author

Email addresses: swen@ysu.edu.cn (Shuhuan Wen), dimpledhan@qq.com (Xiaohan Lv), hak-keung.lam@kcl.ac.uk (Hak Keung Lam), fskfsk399@163.com (Shaokang Fan), 309926190@qq.com (Xiao Yuan), ming.1.chen@kcl.ac.uk (Ming Chen)

1. Introduction

Autonomy in the complex environment is very important to mobile robot, which requires the robot can navigate autonomously, avoid obstacles and mapping. Localization and mapping have been well-studied in the past decades based on simultaneous localization and mapping (SLAM)[1]. However, a path planning algorithm integrated with SLAM to form an all-encompassing framework is still in its early years.

State-of-the-art planning algorithms mostly assume that accurate localization is available and the environment information is at least partially known[2]. But real-world environment is always unknown and complex, which is an interesting challenge for robust autonomy. The obstacle avoidance algorithm can make the robots move autonomously and remove the restrictions of space on robots. So the path planned by the robot in the complex environment can greatly affect the performance of SLAM. Thus, a comprehensive framework integrating localization, mapping and planning is demanded for real-world mobile robot. An active SLAM framework that integrates path planning algorithm with SLAM is proposed to construct environmental map and realize obstacle avoidance in unknown and complex environment[3, 4].

In this paper, we propose an improved active SLAM algorithm based on our previous framework[3] and Monodepth depth image recognition method. The training of depth map prediction and path planning based on deep reinforcement learning (DRL) can achieve better performance of obstacle avoidance in the complex environment. DRL is used to plan the path of avoidance obstacles, then the robot can successfully navigate and complete the map construction of the environment. In our experiments, the mobile robot Turtlebot is equipped with a Kinect sensor and a laser. Laser is used to build the map of the environment, and Kinect is used to perceive the external environment and recognize the obstacles. The proposed algorithm can improve the prediction speed of the depth image and optimize the walking path of the robot. The contributions of our work are as follows:

- 1) Binocular stereo image for the obstacles by using Monodepth method is obtained and its prediction speed is faster than Fully Convolutional Residual Networks (FCRN).
- 2) Probability Dueling DQN (PD3QN) is used to optimize the path of the robot,
35 which can reach the destination more quickly than Dueling DQN.
- 3) PD3QN algorithm is combined with FastSLAM to achieve the autonomous navigation and map the environment in the complex environment with different numbers and sizes of the static and dynamic obstacles.

2. Related Work

40 Traditional SLAM needs to pre-set the trajectory of mobile robot and avoid obstacles in order to achieve the purpose of building an environmental map, so it is called passive SLAM. However, in practical applications, most of the environment is unknown. Before entering the environment, the robot cannot plan a moving path in advance. Active SLAM is a good solution to establish
45 an accurate and complete environmental map in unknown environment[4]. Active SLAM originates from the work of Bajcsy[5] and finding control actions to reduce the uncertainty. Sim and Roy[6] only optimized the trajectory of the SLAM covariance matrix, but did not include a measure of the path length in the objective function. Ref.[7] included path lengths and reduces uncertainty
50 while conducting effective exploration. Galceran[8] proposed a survey planning method to minimize the attitude uncertainty of autonomous underwater vehicle (AUV). Valencia[9] proposed an active SLAM framework to assess the uncertainty of the posture map of the revisiting actions. However, when deciding on the best action, only a few possible choices are listed in the expression. And their
55 approach is focused on the details of bathymetric surveys and is not integrated with the SLAM framework. Sampling-based planning, such as RRTS[10] and HFPRM[11] gradually improve the shortest-distance path planning while maximizing the optimality. These methods provide local optimization, but have not yet expanded to a real-world robotic system. In this paper, the path planning

60 method and SLAM framework are combined to form an active SLAM method, which ensures that the robot can navigate autonomously and avoid obstacles effectively in the unknown and complex environments, and successfully complete the construction of environmental map.

In this paper, deep reinforcement learning (DRL) is used to plan a path integrated with FastSLAM. Path planning means that robots can autonomously
65 avoid obstacles and reach the target point[11]. According to the state of obstacles, obstacles are divided into static obstacles and dynamic obstacles[12]. Dynamic Window Approach[13, 14], Visibility Graph[15], A* Search Algorithms[16, 17, 18] and D* algorithm[19] are all classic static obstacle avoidance methods. Dynamic Window Approach is suitable for the operation of high-speed mobile
70 robots, which can directly calculate the robot rotation and translation control commands of the robots in the velocity space. Visibility Graph realizes obstacle avoidance path planning by finding the barrier-free connection between the starting point and the vertices of each obstacle, so it needs a very accurate
75 description of obstacles in the environment. A* Search algorithm is a path planning obstacle avoidance algorithm based on Estimation Functions[16] to search for optimal path. Therefore, the quality of Estimation Functions has a great influence on its planning path. In the complex environment, the traditional obstacle avoidance algorithm cannot obtain acceptable performance. At
80 present, there are few studies on dynamic obstacle avoidance. D* algorithm is an early developed and mature dynamic obstacle avoidance algorithm[19]. It is improved based on A* algorithm and can be applied to dynamic obstacle avoidance. When the environment changes, although it can avoid obstacles to reach the destination, the planned path is not guaranteed to be optimal.

85 With the development of artificial intelligence, more and more obstacle avoidance algorithms that can adapt to the complex environments with static and dynamic obstacles have been proposed, such as Genetic Algorithm[20], Fuzzy Control[21], Reinforcement Learning[22], Deep Reinforcement Learning[23, 24]. The advantage of obstacle avoidance algorithm based on Genetic Algorithm
90 is that it is simple and easy to implement, and the disadvantage is that it tends

to appear premature convergence. Fuzzy Control has better applicability and fault tolerance, and the disadvantage is that the fuzzy rules and membership functions need to be determined by experience. Reinforcement Learning has better generalization ability, and the disadvantage is that the slow learning speed leads to lower convergence speed. Deep Reinforcement Learning has good adaptability and can be applied to wide range of environments, but it needs a lot of training. In this paper, three experimental scenes are designed to test the proposed algorithm, include three static obstacles in scene 1, four static obstacles in scene 2, and two static obstacles and one dynamic obstacle in scene 3. The experiments show that the image predictive method based on Monodepth decreases the predictive time comparing with the traditional Fully Convolutional Residual Networks (FCRN) method. The proposed Probability Dueling DQN (PD3QN) can plan the path and make the robot reach the destination using less time comparing with single D3QN under environment with different numbers of static and dynamic obstacles.

The rest of this paper is organized as follows. Section II gives the review of related works in active SLAM. Section III introduces the method of Depth map prediction based on Monodepth. Section IV shows the framework of Active FastSLAM based on Probabilistic Dueling DQN. Experimental results are presented in Section V, and conclusions are shown in Section VI.

3. Depth map prediction based on Monodepth

Learning-based approach has achieved good results in single image depth estimation. Most existing methods regard depth prediction as a supervised regression problem, so a large number of corresponding depth data of the ground truth is trained. How to record the depth data with high quality in various environments is a challenging task.

Monodepth is a depth image prediction method for single image based on unsupervised deep neural network. Depth is estimated by inferring the difference of the left image distorted to match the right. The two differences (from

120 left to right and from right to left) can be inferred simultaneously, then better
 depth can be obtained by using only the left input image and forcing them to
 be consistent with each other. The network uses bilinear sampler and gener-
 ates predictive image by reverse mapping, thus obtaining a fully differentiable
 image generation model. To align the output disparity map with the left input
 125 image, the network must sample from the right image[6]. The network can be
 trained to generate a left view by sampling from the right image to create a
 left view aligned disparity map. To solve this problem, a trained network is
 used to predict the disparity maps of two views by sampling the input images.
 A single left image is used as input to the convolution layer, while the right
 130 image is only used in training. By using this new left-right consistency cost to
 enhance consistency between the two disparity maps, more accurate results can
 be obtained. Although the network accepts only a single image as input, two
 disparity maps (from left to right and right to left) are predicted on each output
 scale.

135 In the training process, network learning generates images by sampling pixels
 from relative stereo images. The image model uses image samplers from the
 Spatial Transformer Network (STN) to sample input images using disparity
 maps. STN adopts bilinear sampling, and the output pixel is a weighted sum
 of four input pixels. The bilinear sampler is locally completely differentiable
 140 and integrated into the full convolution architecture without any simplification
 or approximation of the cost function. In order to generate a more accurate
 disparity map, the trained the network predicts the parallax of the left and right
 images and only the left view is used as the input of the network convolution
 portion. In order to ensure consistency, the consistency penalty of the left and
 145 right parallax is introduced to the model. It attempts to make the parallax of
 the left view equal to the projected parallax of the right view. In the test, the
 network predicts the parallax of the left image so that it has the same resolution
 as the input image.

Table 1: Time (seconds) required for FCRN and Monodeth to predict depth maps

Experimental Scenario	FCRN	Monodepth
1	1.518595	1.262896
2	1.538330	1.263206
3	1.533961	1.272562
4	1.555996	1.268395
5	1.542547	1.266421

For autonomous navigation and obstacle avoidance, it is important to identify obstacles more accurately and clearly with the least time. In this paper, Monodepth prediction method is used to identify obstacles, and binocular stereo image replaces the depth data used in the FCRN training, which can make the convolution Neural networks perform single image depth estimation in the absence of ground truth depth data. The network is trained to generate parallax image through hyperpolar geometric constraints and image reconstruction Loss. The experiment results of the depth image using Monodepth method for the 5 experimental scenarios are shown in Figure 1. From left to right, Figure 1 are the RGB image, the FCRN depth prediction image, and the Monodepth depth prediction image for the experimental scene. The time using FCRN and Monodeth method for the experimental scenarios above is shown in Table 1, which shows that the time of predicting the depth image using FCRN is about 1.5 seconds, and the time of predicting the depth image using Monodepth is about 1.2 seconds. Monodepth shortens the prediction time and improves the prediction effect comparing with FCRN.

Figure 1: The results of FCRN depth prediction images and Monodeth depth prediction images

165 4. Active FastSLAM based on Probabilistic Dueling DQN

In traditional Q-learning, when the state and action space are discrete and the dimensions are not high, the Q-table can be used to store the Q-value of each state-action pair. When the state and action space are of high-dimensional and continuous, it is unrealistic to use Q-table. Generally, the updating problem of Q-table is transformed into a function fitting problem. Similar states lead to similar output actions. By updating parameters, Q-function approximates the optimal Q-value. For the value of the state-action pair (s,a) , the formula with parameters θ is as follows.

$$Q(s, a; \theta) \approx Q'(s, a) \quad (1)$$

Deep neural networks can automatically extract complex features, so it is suitable to use deep neural network when facing high-dimensional continuous state[7]. Dueling DQN improves the network structure. It uses the model structure to express value function in more details, which makes the model obtain better performance and reduces the overestimation of Q-value.

Dueling-DQN constructs two completely connected hierarchical flows, which is different from traditional DQN structure. One is state value estimation, the other is motion advantage function with independent of state. The state value flow and the action advantage flow are calculated separately, and finally combined together to calculate the Q-value. The calculation of Q and state value V is updated simultaneously. The advantage function of different actions is calculated in the same state, and the state value flow will be recalculated, which means that the updating efficiency of the value flow is higher.

The learning module of the basic convolution feature is shown in Figure 2[25]. One of the main benefits of module decomposition is that it generalizes learning between behaviors without changing the reinforcement learning algorithm. Duel networks should also be understood as a single Q network, with two rather than one flow. Duel networks can know if the states are valuable or not without knowing the effect of each action in each state.

Figure 2: Dueling DQN network is divided into state value flow and action advantage flow

The value function Q is decomposed into the state value function (V) and the dominant function (A). For a policy π , the expression is as follows.

$$Q^\pi(s, a) = V^\pi(s) + A^\pi(s, a) \quad (2)$$

195 The advantage function shows the difference between current behavior and average performance. If it is better than the average value, the advantage function is positive. If it is worse than the average value, the advantage function is negative. We add a limitation to the advantage function. We know that the expectation of the advantage function is zero. The value function Q is described
200 as follows, where $A(s, a')$ is an $|A|$ -dimensional vector.

$$Q^\pi(s, a) = V^\pi(s) + (A^\pi(s, a) - \frac{1}{|A|} \sum_{a'} A(s, a')) \quad (3)$$

By subtracting the average from each value, it can guarantee a constraint with the expected value of 0, which increases the stability of the collected output. There is no need to train a network when ordinal values are used in some cases. The output value of the function is given by an explicit way, then the function is
205 explicitly updated each time[8]. This can increase the update frequency of the

function. For single-output Q-network, there is some ambiguous for its update, which makes it user-friendly for network training.

Reinforcement learning is a very effective way to perform control decisions[9, 10, 11]. Deep learning adopts an end-to-end approach for high-dimensional information input such as depth image and color image[12, 13, 14]. Deep reinforcement learning algorithm has a good effect on path planning of mobile robots. Since deep learning needs to be trained in the simulation environment before it is applied to the actual environment, the training simulation environment is built in the Gazebo simulator. Because Dueling DQN (D3QN) model is superior to Deep Double Q network (DDQN) and DQN, this paper adopts Dueling DQN model[15] to plan the path of the mobile robot. When a robot walks in an unknown environment, RGB images can be transformed into RGB-D images by using visual information, and the walking space and obstacle space can be inferred. The robot will adopt D3QN algorithm to avoid obstacles and reach the target position.

For reinforcement learning, when an agent performs every action a_t , it will receive immediate reward r_t for the action. Different actions have different reward values, and the size of the immediate reward value determines the quality of the algorithm[16]. Therefore, each action selection is important for the robot to take the optimal path in the end[17]. For traditional Q-learning algorithm, the agent randomly chooses the action to operate in the action space every time, that is, the agent will choose all possible actions with equal probability in any state. However, if each action in action space is given equal selection probability, the path will not be the optimal in the path planning. In some states, other actions may be chosen instead of the optimal action. Therefore, each action in the action space is given according to different chosen probabilities, which can optimize the path. When the environmental state is s , the agent makes a probabilistic choice for all possible actions in the action space. The probability of action a_i is $P(a_i | s)$, and k is a constant and $k > 0$. The probability of the

235 chosen action is expressed in Eq(4).

$$P(a_i | s) = \frac{k^{Q(s,a_i)}}{\sum_j k^{Q(s,a_j)}} \quad (4)$$

If the Q-value of an action is higher, the performing action probability of the agent is higher. Therefore, the agent will choose the action with higher reward value in each state, which will make the path better. k value reflects the importance of Q-value action selection for the agent. The action with higher Q-value than the average level will have a larger k value and get a higher probability
240 of action selection. In this way, agents tend to select actions with higher Q-values to perform in order to get the maximum reward value. The smaller k value also makes the action with lower Q-value have relatively higher selection probability, and will not make the selection probability of an action become zero.

245 For the classical Dueling DQN obstacle avoidance algorithm, each action is given different probability according to (1), which makes the robot choose the action with higher reward value when performing the action. In this paper, the Dueling DQN algorithm with probabilistic selection action is called probability Dueling DQN (PD3QN). Dueling DQN with probability choice will make the planned
250 path shorter and the walking steps less than classical Dueling DQN algorithm. In this paper, FastSLAM is used to construct the map of the environment, and the improved Probability Dueling DQN (PD3QN) algorithm is used to plan the path of the obstacle avoidance, which forms an active SLAM framework. The framework of the proposed algorithm is shown in Figure 3.

Figure 3: Framework of active SLAM algorithm based on Probabilistic Dueling DQN

255 5. Experiments

In order to verify the validness of the improved algorithm, three groups of experiments are designed in this paper. The first group is set up as three static

obstacles in the environment, the second group is set up as four static obstacles in the environment, and the third group is set up as two static obstacles and one dynamic obstacle (walking person in this paper) in the environment. The experimental results of Probability Dueling DQN (PD3QN) algorithm based on Monodepth image prediction are compared with Dueling DQN algorithm based on FCRN image prediction. Different static and dynamic obstacles are used to test the performance of the proposed algorithm. In order to compare the path length using the two algorithms, the target point is set and the experimental scene in the same environment is adopted.

The experimental scene of three columnar static obstacles is shown in Figure 4a). Figure 4b) is the environmental map built by active FastSLAM based on Dueling DQN obstacle avoidance algorithm, and FCRN is used to predict depth image. In Figure 4b), the trajectory of the robot is a red curve, the black point is an obstacle, and the black solid circle is a robot. Figure 4c) is the planned path of the robot in Figure 4b). Figure 4d) is the environment map using the proposed algorithm in this paper. Figure 4e) is the walking trajectory of the robot in Figure 4d). In Figure 4c) and Figure 4e), the trajectory of the robot is an orange curve. The black four-point star is the obstacle, the yellow solid circle is the starting point, the green solid circle is the end point. The robot starts its navigation from the starting position. The robot can recognize and avoid the three static obstacles and reach the target position. At the same time, the map of the experimental environment is built. The walking steps of the robot using Dueling DQN based on FCRN and Probability Dueling DQN based on Monodepth active FastSLAM are shown in Table 2.

The experimental scene of three columnar static obstacles and one box-shaped static obstacle is shown in Figure 5. Figure 5 shows that the robot still can avoid the four obstacles and build the environmental map successfully even if the numbers of the obstacles increases. And the walking steps of the robot are shown in Table 2.

The experimental scene with a dynamic obstacle (walking person) and two columnar static obstacles is shown in Figure 6. The environment map and the

trajectory of the robot are shown in Figure 6b) and Figure 6c). In Figure 6c) and Figure 6e), the blue solid triangle is the trajectory of the dynamic obstacle, and the pink curve in the orange curve is the new trajectory when the robot encounters the dynamic obstacle. The robot can avoid the two static obstacles during its navigation. When the robot recognizes the obstacle 1 (a walking person), it can also avoid the moving obstacle and reach the target position, at the same time, an environmental map is built. The walking steps of the robot are also shown in Table 2. Table 2 shows that the walking steps using Probability Dueling DQN based on Monodepth active FastSLAM is less than Dueling DQN based on FCRN active FastSLAM. Due to the limited experimental space, the robot does not have more space to walk. If the robot is set to a far target point, the robot will have more obvious differences in the number of steps.

Experimental results show that the Monodepth method can better predict depth images comparing with FCRN method in the same environment, and the prediction time is shorter. The proposed active FastSLAM algorithm in this paper also makes the robot walk in shorter path, improves the efficiency of path planning algorithm and builds the environment map.

Table 2: Walking Steps using Dueling DQN and Probabilistic Dueling DQN avoidance obstacles algorithm in Three Groups of Experiments

Experiment	Obstacles	Dueling DQN	Probabilistic Dueling DQN
1	Three Static Obstacles	52	48
2	Four Static Obstacles	56	50
3	Two Static Obstacles and One Dynamic Obstacle	61	56

(a) Experimental scenario

(b) Environmental map built using Dueling DQN obstacle avoidance based on FCRN active FastSLAM

(c) Robot trajectory using Dueling DQN obstacle avoidance based on FCRN active FastSLAM

(d) Environmental map built using the proposed algorithms in this paper

(e) Robot trajectory of the robot using the proposed algorithms in this paper

Figure 4: Experimental scene, environment map and robot trajectory of three static obstacles

(a) Experimental scenario

(b) Environmental map built using Dueling DQN obstacle avoidance based on FCRN active FastSLAM

(c) Robot trajectory using Dueling DQN obstacle avoidance based on FCRN active FastSLAM

(d) Environmental map built using the proposed algorithms in this paper

(e) Robot trajectory using the proposed algorithms in this paper

Figure 5: Experimental scene, environment map and robot trajectory of four static obstacles

(a) Experimental scenario

(b) Environmental map built using Dueling DQN obstacle avoidance based on FCRN active FastSLAM

(c) Robot trajectory using Dueling DQN obstacle avoidance based on FCRN active FastSLAM

(d) Environmental map built using the proposed algorithms in this paper

(e) Robot trajectory using the proposed algorithms in this paper

Figure 6: Experimental scene, environment map and robot trajectory of two static obstacles and one dynamic obstacle

6. Conclusion

In this paper, we proposed a comprehensive active FastSLAM framework in the complex environment. We presented a path planning algorithm integrated with visual SLAM in order to find an optimal path and built the environment map for a robot exploring a complex environment with different numbers of the static and dynamic obstacles. Monodepth method was used to predict the depth image. We combined Probability Dueling DQN path planning approach for efficiently reaching the destination with a FastSLAM framework for building the environment map. The experimental results showed that our framework offers many benefits over other representative active SLAM method: fast and accurate depth image prediction, short steps of planning paths and consistent good performance over our experimental trials. This type of active SLAM system is useful to accomplish robust, autonomous navigation in the complex environments.

References

- [1] E. Colle, S. Galerne, A robust set approach for mobile robot localization in ambient environment, *Autonomous Robots* 3 (2018) 1–17.
- [2] S. Wen, W. Zheng, J. Zhu, X. Li, S. Chen, Elman fuzzy adaptive control for obstacle avoidance of mobile robots using hybrid force/position incorporation, *IEEE Transactions on Systems Man Cybernetics Part C Applications Reviews* 42 (2012) 603–608.
- [3] S. Wen, X. Chen, C. Ma, H. K. Lam, S. Hua, The Q-Learning obstacle avoidance algorithm based on EKF-SLAM for NAO autonomous walking under unknown environments, *Robotics Autonomous Systems* 72 (2015) 29–36.
- [4] M. L. Rodriguez-Arevalo, J. Neira, J. A. Castellanos, On the importance of uncertainty representation in active SLAM, *IEEE Transactions on Robotics* 34 (2018) 829–834.

- [5] R. Bajcsy, Active perception, *Proceedings of the IEEE* 76 (1988) 966–1005.
- 335 [6] R. Sim, N. Roy, Global A-Optimal robot exploration in SLAM, in: *Proceedings of IEEE International Conference on Robotics and Automation*, 2006. doi:10.1109/ROBOT.2005.1570193.
- [7] H. H. Gonzalez-Banos, J.-C. Latombe, Navigation strategies for exploring indoor environments, *The International Journal of Robotics Research* 21 (2002) 829–848.
- 340 [8] E. Galceran, S. Nagappa, M. Carreras, P. Ridao, A. Palomer, Uncertainty-driven survey path planning for bathymetric mapping, in: *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2013, pp. 6006–6012. doi:10.1109/IRoS.2013.6697228.
- 345 [9] R. Valencia, M. Morta, J. Andrade-Cetto, J. M. Porta, Planning reliable paths with pose SLAM, *IEEE Transactions on Robotics* 29 (2013) 1050–1059.
- [10] S. M. LaValle, J. Kuffner, Randomized kinodynamic planning, *International Journal of Robotics Research* 15 (1999) 378–400.
- 350 [11] K. Wei, B. Ren, A method on dynamic path planning for robotic manipulator autonomous obstacle avoidance based on an improved RRT algorithm, *Sensors* 18 (2018) 571.
- [12] S. Samavati, M. Zarei, M. T. Masouleh, An optimal motion planning and obstacle avoidance algorithm based on the finite time velocity obstacle approach, in: *Proceedings of Artificial Intelligence Signal Processing Conference*, 2018, pp. 250–255. doi:10.1109/AISP.2017.8324091.
- 355 [13] A. Ozdemir, V. Sezer, A hybrid obstacle avoidance method: Follow the gap with dynamic window approach, in: *Proceedings of IEEE International Conference on Robotic Computing*, 2017, pp. 257–262. doi:10.1109/IRC.2017.25.
- 360

- [14] D. Lee, J. Jeong, H. K. Yong, B. P. Jin, An improved artificial potential field method with a new point of attractive force for a mobile robot, in: Proceedings of International Conference on Robotics Automation Engineering, 2017, pp. 63–67. doi:10.1109/ICRAE.2017.8291354.
- 365 [15] D. Dakopoulos, N. G. Bourbakis, Wearable obstacle avoidance electronic travel aids for blind: A survey, IEEE Transactions on Systems Man Cybernetics Part C 40 (2009) 25–35.
- [16] D. J. Seo, J. Kim, Development of autonomous navigation system for an indoor service robot application, in: Proceedings of International Conference on Control, 2014, pp. 204–206. doi:10.1109/ICCAS.2013.6703894.
- 370 [17] J. Yao, C. Lin, X. Xie, A. J. W. Wang, C.-C. Hung, Path planning for virtual human motion using improved A* star algorithm, in: Seventh International Conference on Information Technology: New Generations, 2010, pp. 1154–1158. doi:10.1109/ITNG.2010.53.
- 375 [18] G. Sahin, M. Balcilar, E. Uslu, S. Yavuz, M. Amasyali, Obstacle avoidance with vector field histogram algorithm for search and rescue robots, in: Proceedings of Signal Processing Communications Applications Conference, 2014, pp. 766–769. doi:10.1109/SIU.2014.6830342.
- [19] Z.-K. Chen, S. Li, Application of matching optical flow method in obstacle avoidance of mobile, Journal of Discrete Mathematical Sciences and Cryptography 21 (2018) 1119–1129.
- 380 [20] Q. Zhang, J. Ma, W. Xie, L. Ma, Improved ant colony algorithm-based path planning for mobile robot, Journal of Northeastern University 34 (2013) 1521–1524.
- 385 [21] M. Li, H. Cao, X. Song, Y. Huang, Shared control driver assistance system based on driving intention and situation assessment, IEEE Transactions on Industrial Informatics 14 (2018) 4982–4994.

- 390 [22] Y. Wang, H. He, C. Sun, Learning to navigate through complex dynamic environment with modular deep reinforcement learning, *IEEE Transactions on Games* 10 (2018) 400–412.
- [23] Z. Ning, R. Chen, Y. Xia, H. Jie, G. Wen, Neural network-based reconfiguration control for spacecraft formation in obstacle environments, *International Journal of Robust Nonlinear Control* 28 (2018) 2442–2456.
- 395 [24] C. Yin, W. Zhang, Concise deep reinforcement learning obstacle avoidance for underactuated unmanned marine vessels, *Neurocomputing* 272 (2018) 63–73.
- [25] Z. Wang, T. Schaul, M. Hessel, H. van Hasselt, M. Lanctot, N. de Frcitas, Dueling network architectures for deep reinforcement learning, volume 4, 2016, pp. 2939–2947.