

Underwater Target Detection Based on Reinforcement Learning and Ant Colony Optimization

WANG Xinhua^{1), *}, ZHU Yungang²⁾, LI Dayu³⁾, and ZHANG Guang³⁾

1) School of Computer Science, Northeast Electric Power University, Jilin 132012, China

2) Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, College of Computer Science and Technology, Jilin University, Changchun 130012, China

3) State Key Laboratory of Applied Optics, Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China

(Received January 19, 2021; revised June 30, 2021; accepted September 10, 2021)

© Ocean University of China, Science Press and Springer-Verlag GmbH Germany 2022

Abstract Underwater optical imaging produces images with high resolution and abundant information and hence has outstanding advantages in short-distance underwater target detection. However, low-light and high-noise scenarios pose great challenges in underwater image and video analyses. To improve the accuracy and anti-noise performance of underwater target image edge detection, an underwater target edge detection method based on ant colony optimization and reinforcement learning is proposed in this paper. First, the reinforcement learning concept is integrated into artificial ants' movements, and a variable radius sensing strategy is proposed to calculate the transition probability of each pixel. These methods aim to avoid undetection and misdetection of some pixels in image edges. Second, a double-population ant colony strategy is proposed, where the search process takes into account global search and local search abilities. Experimental results show that the algorithm can effectively extract the contour information of underwater targets and keep the image texture well and also has ideal anti-interference performance.

Key words ant colony optimization; reinforcement learning; underwater target; edge detection

1 Introduction

In recent years, underwater optical imaging equipment has rapidly developed, and underwater target detection technology has also been widely used (Lin *et al.*, 2019). Such technology involves the laying of submarine optical cables (Fatan *et al.*, 2016), the establishment and maintenance of underwater oil platforms (Bonin-Font *et al.*, 2015), the salvage of submarine sinking ships (Liu *et al.*, 2019), and the research of marine ecosystems (Li *et al.*, 2020). Underwater optical imaging produces images with high resolution and abundant information and hence has outstanding advantages in the short-distance underwater target detection task. However, due to the influence of water light absorption and scattering, images captured by the underwater optical imaging system often encounter many problems, such as increased noise interference, blurred texture features, low contrast, and color distortion (Wang *et al.*, 2019). Therefore, the underwater target detection task faces many challenges. How to detect the underwater target accurately, quickly, and stably in a complex underwater scene with poor image visibility has become an urgent problem that needs to be solved. In recent years, domestic

and foreign research institutions and scholars have conducted considerable research on the underwater target detection method, which can be divided into underwater target detection based on traditional features and underwater target detection based on deep learning networks. The traditional underwater target detection method based on feature descriptors describes underwater targets. The commonly used underwater image features include color, shape, and texture features, which is a simple method and has good real-time performance (Chuang *et al.*, 2016). However, they are also affected by changes in the target size, rotation, occlusion, shooting angle, and species category. With the development of the graphics processing unit (GPU) and other hardware systems, target detection technology based on deep learning has rapidly developed. In the computer vision task, a deep learning network extracts information layer by layer from pixel-level raw data to abstract semantic concepts, which makes it have outstanding advantages in extracting global features and context information of images. Even in the case of occlusions or small object size, it can perform successful detection (Sun *et al.*, 2018). However, due to the complex structure of deep neural networks, they need to adjust several parameters when applied to a specific environment, resulting in a decline in the algorithm efficiency.

In the fields of ecological monitoring of the marine

* Corresponding author. E-mail: wangxh@neepu.edu.cn

biological population, the edge contour of a target image is an important appearance feature used to distinguish marine biology. By detecting the edge of a target image, not only the feature information of the target can be effectively extracted, but also the computational complexity and amount of data processing can be greatly reduced. As a metaheuristic algorithm, the ant colony optimization (ACO) algorithm is insensitive to noise. Accordingly, this study attempts to apply the ACO algorithm to underwater target image edge detection to improve the image edge detection accuracy and anti-noise ability. Our contributions are as follows:

1) The reinforcement learning idea is integrated into the artificial ant movement, and a variable sensing radius strategy is proposed to calculate the transition probability of each pixel to avoid undetection or misdetection of some pixels in an image edge that cannot be detected or even mistakenly detected.

2) The double-population strategy is proposed to control the movement direction of ants, the first population focuses on the global search, and the second population focuses on the local search. Thus, the search process takes into account global search and local search abilities.

3) The proposed method can achieve high accuracy in practical datasets in the real world.

2 Principle of the Algorithm

2.1 ACO

The ACO algorithm is a heuristic search algorithm that simulates ant foraging. As a probabilistic group intelligent search technology, ACO has been widely used in many fields, such as data mining (Gao *et al.*, 2013), cluster analysis (Deng *et al.*, 2019), and image processing (Tian *et al.*, 2008). The application of the ACO algorithm in image edge detection was first proposed by Nezamabadi-pour *et al.* (2006). The flow of the image edge detection algorithm based on ACO is shown in Fig.1.

In Fig.1, the basic principle of the ACO algorithm is that ants will deposit pheromones along the path they travel during foraging, and other ants can detect pheromones to guide their movement (Lu and Chen, 2008). During the initialization of the ACO algorithm, a predetermined number of artificial ants are placed in the search space. The movement of each ant is based on the transition probability, which indicates the probability of the ant moving from one unit to another in the search space. The value of the transition probability depends on the pheromone and heuristic information of the unit. After the ants move, the pheromone is updated. When the predefined number of iterations is reached, the search is terminated. Then, most ant travel routes that contain more pheromones are selected as the best solution.

2.2 Reinforcement Learning

The basic idea of reinforcement learning is, if a certain action of the system can bring positive returns, then the future trend of the action will be stronger; otherwise, the trend of the action will be weakened (Montague, 1999).

Based on the reinforcement learning model, a reinforcement learning system includes not only the agent and environment but also the four basic elements, namely, strategy, value function, return function, and environment model. Because the state transition probability function and reward function in the environmental model are unknown, the agent can only choose a strategy based on the rewards obtained from each trial. Therefore, a function for strategy selection needs to be constructed between the strategy and instantaneous reward (Mnih *et al.*, 2015).

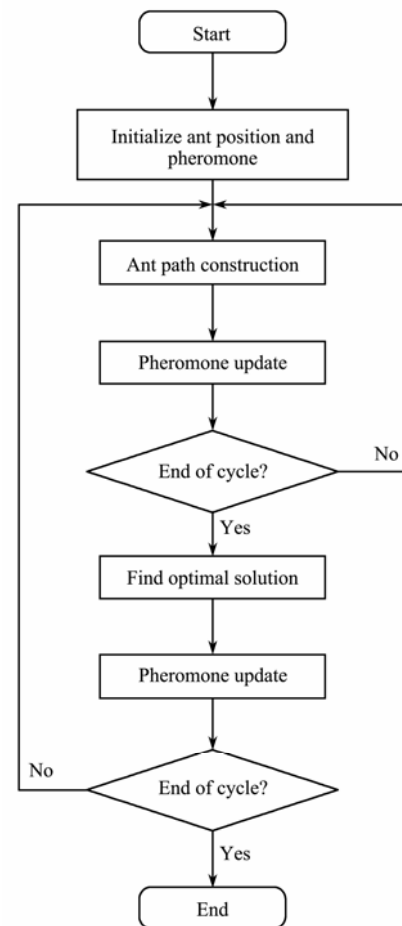


Fig.1 Flow of the image edge detection algorithm based on ACO.

Reinforcement learning is a heuristic learning strategy based on machine learning theory. The problem to be solved is how to choose the action that can achieve the optimal goal through learning (Silver *et al.*, 2018). First, you need to build an agent that can sense the surrounding environment and complete the mapping from environment to action. Reinforcement learning regards learning as a process of continuous exploration. The standard reinforcement learning model is shown in Fig.2.

In Fig.2, the agent accepts the input state s of the environment and outputs the corresponding action a through the internal learning inference mechanism. Under action a , the environment becomes a new State s' . At the same time, an enhanced signal r for a reward or punishment is generated and fed back to the agent. The agent selects the next

operation based on the feedback information and the current state of the environment (Vinyals *et al.*, 2019).

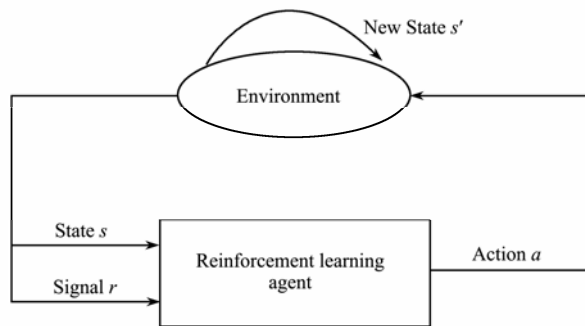


Fig.2 Reinforcement learning model.

3 Related Work

In recent years, many scholars have studied the application of the ACO algorithm in image edge detection. Most of these methods improve the algorithm from two aspects: the improvement based on the algorithm itself and the fusion with other algorithms. Singh *et al.* (2019) proposed an ACO method based on bioinspired technology to identify the edges of ships in the ocean. To reduce the time complexity, a triangular fuzzy membership function is used. The results of the proposed work confirm the clear edges of small and partial objects. Kheirinejad *et al.* (2018) applied the max-min ACO method to detect image edges. Moreover, a new heuristic information function (HIF) is proposed, namely, group-based HIF, to determine nodes that ants visit around their place. The proposed HIF exploits the difference between the intensity of two groups of nodes instead of two single ones. The simulation result sections show that the robustness of the proposed edge detection algorithm is more than that of the previous algorithms. Nasution *et al.* (2017) proposed improvements to the edge detection method with the approach graph with the ACO algorithm. The repairs may be performed to thicken the edge and connect the cutoff edges. The ACO method can be used as a method for optimizing operator Robinson by improving the image result of the Robinson detection average by 16.77% as compared with that of the classic operator Robinson. Rafsanjani and Varzaneh (2015) proposed an approach based on the distribution of ants on an image. Here ants try to find possible edges using a state transition function. Experimental results show that the proposed method is less sensitive to Gaussian noise compared to standard edge detectors and gives finer details and thinner edges compared to earlier ant-based approaches. Liu and Fang (2015) proposed an edge detection method based on the ACO algorithm. This method uses a new heuristic function and a user-defined threshold during the pheromone update process and provides a suitable set of parameters. Experimental results show that the method can also show good effectiveness in the presence of noise. Dawson and Stewart (2014) presented the first parallel implementation of an ACO-based edge detection algorithm on the GPU using NVIDIA CUDA. By exploiting the mas-

sively parallel nature of the GPU, we can execute significantly more ants per ACO iteration, allowing us to reduce the total number of iterations required to create an edge map. We hope that reducing the execution time of an ACO-based implementation of edge detection will increase its viability in image processing and computer vision. For the image edge detection problem in complex scenes, most ACO algorithms use a fixed number of neighborhood pixels to calculate the pixel gradient, which we call a fixed sensing radius, and then heuristic information that calculates the transition probability of each artificial ant at each pixel. However, this strategy may result in the undetection and misdetection of some pixels at the edge of the image. In other words, it may lose some important edges or detect worthless edges. In addition, the movement of artificial ants for complex scene images will produce precocity, resulting in only local optimal solutions, not global optimal solutions.

4 Proposed Method

To solve the abovementioned algorithm problems, this paper proposes an image edge detection algorithm that combines reinforcement learning and ACO. This method is different from the traditional method that uses a fixed and constant number of neighborhood pixels to calculate the gradient. By integrating multiagent reinforcement learning with the movement of artificial ants, the number of neighborhood pixels used to calculate the gradient of each pixel can be adaptively determined. In addition, the dual-population strategy and adaptive parameters are introduced to control the movement direction of the artificial ant so as to prevent local optimization.

4.1 Parameter Initialization of the ACO Distribution

The image of $m \times n$ pixels is mapped into a two-dimensional grid, and each pixel of the image corresponds to a cell of the grid. The number of ants is set to $N = \sqrt{m \times n}$, and N ants are randomly distributed on the pixels.

The pheromone matrix is constructed by moving artificial ants on the image grid, and the image edge detection is based on the pheromone matrix. Artificial ants deposit pheromones by constantly moving the pixels in the image grid. The initial value of pheromones is set to a random value. When the threshold of the convergence time or a given number of times is reached, pixels with higher pheromones will have a greater probability of belonging to the edge of the image.

4.2 Reinforcement Learning Transfer Probability Strategy

Reinforcement learning is a field of machine learning inspired by behavioral psychology. It studies how agents should take actions in the environment to maximize the cumulative rewards (Lewis and vrabie, 2009). When calculating the gradient information of pixels in the proof of artificial ant transfer, the selection of sensing radius will affect the effect of edge detection. As mentioned before,

most methods adopt the strategy of having the same sensing radius in all pixels throughout the search process. This condition affects the detection effect, which may make the algorithm lose some important edges or detect worthless edges (Kober *et al.*, 2013). Considering the actual situation around each pixel, using different sensing radii for different pixels is a promising optimization approach.

We propose a new variable radius sensing strategy to calculate the gradient of each pixel. Reinforcement learning is integrated into the artificial ant motion, and then the perception radius of each pixel is obtained. In the reinforcement learning strategy, the state of the ant is its position in the $m \times n$ image grid, that is, $\text{State} = (r, s)$, and (r, s) represents the pixel coordinates. The action of an ant is a binary $\text{Action} = (a, k)$, which means that an action contains two subactions. The first action a is the movement from the current position to the adjacent position in the image grid, and the second action k is the perception radius. After ants choose an action in a state, they will be rewarded. The $Q(\text{State}, \text{Action})$ function, that is, $Q((r, s), a, k)$, is the maximum reward obtained by executing action a in state (r, s) and selecting perception radius k . The Q function is initially set to a random positive value. After each artificial ant moves, its updating formula is formulated as follows:

$$Q(\text{State}, \text{Action}) = Q((r, s), a, k) \leftarrow \text{reward} + \gamma \max_{a', k'} Q((r, s)', a', k'), \quad (1)$$

where (r, s) is the next state of (r, s) after applying (a, k) action. All ants share the Q function and maintain the Q function.

During the movement of artificial ants, the sensing radius is determined by $P(k|(r, s), a)$. This probability can be calculated by Eq. (2):

$$P(k|(r, s), a) = \frac{P((r, s), a, k)}{P((r, s), a)} \propto P((r, s), a, k). \quad (2)$$

In Eq. (2), the denominator does not include k , so Eq. (3) is proportional to the numerator.

$$P((r, s), a, k) = \frac{Q((r, s), a, k)}{\sum_{a, k} Q((r, s), a, k)}. \quad (3)$$

In Eq. (3), the selection probability $P(k|(r, s), a)$ of the perceptual radius can be obtained by the Q function.

In summary, the new transition probability combining ants and reinforcement learning is defined in Eq. (4):

$$P_{(r, s) \rightarrow (x, y, k)}^i = \begin{cases} \frac{(\tau_{(x, y)})^\alpha (\psi_{(x, y, k)})^\beta (\theta_{(i, x, y)})^\gamma}{\sum_{\substack{(x, y) \\ \in \Omega(r, s)}} (\tau_{(x, y)})^\alpha (\psi_{(x, y, k)})^\beta (\theta_{(i, x, y)})^\gamma}, & \text{if } (x, y) \in \Omega(r, s) \\ 0 & , \text{ otherwise} \end{cases} \quad (4)$$

In Eq. (4), (x, y) represents the pixel coordinates, $\tau_{(x, y)}$ represents the pheromone position (x, y) , $\psi_{(x, y, k)}$ is the heuristic information of the pixel (x, y) , k represents the perceived radius of a pixel (x, y) , and $\theta_{(i, x, y)}$ is the direction factor that controls the movement direction of the ant in the population i . $\Omega(r, s)$ is a collection of neighborhoods containing pixels (r, s) . α is expressed as the $\tau_{(x, y)}$ factor, β is expressed as the $\psi_{(x, y, k)}$ factor, and γ is expressed as $\theta_{(i, x, y)}$.

Heuristic information guides the movement trends of artificial ants. The new heuristic function $\psi_{(x, y, k)}$ is defined as follows:

$$\psi_{(x, y, k)} = P(k|(r, s), a_{(r, s) \rightarrow (x, y)}) \cdot \frac{\ln(\text{grad}_k I_{(x, y)})}{\sum_{i=1}^M \sum_{j=1}^N \ln(\text{grad}_k I_{(x, y)})}, \quad (5)$$

where grad_1 and grad_2 calculation formulas are as follows:

$$\text{grad}_1 I_{x, y} = |I_{x-1, y-1} - I_{x+1, y+1}| + |I_{x-1, y} - I_{x+1, y}| + |I_{x-1, y+1} - I_{x+1, y-1}| + |I_{x, y-1} - I_{x, y+1}|,$$

$$\text{grad}_2 I_{x, y} = |I_{x-2, y-1} - I_{x+2, y+1}| + |I_{x-2, y+1} - I_{x+2, y-1}| + |I_{x-1, y-2} - I_{x+1, y+2}| + |I_{x-1, y-1} - I_{x+1, y+1}| + |I_{x-1, y} - I_{x+1, y}| + |I_{x-1, y+1} - I_{x+1, y-1}| + |I_{x-1, y+2} - I_{x+1, y-2}| + |I_{x, y-1} - I_{x, y+1}|.$$

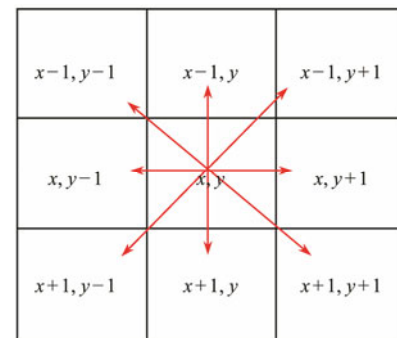


Fig.3 Grad_1 calculation strategy.

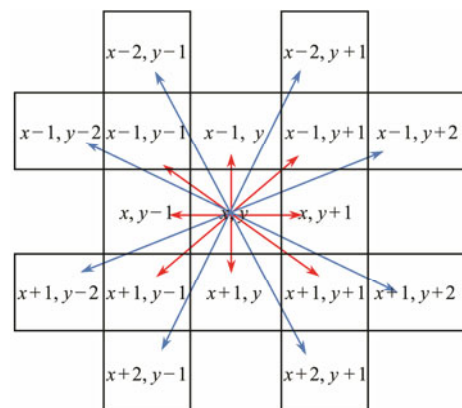


Fig.4 Grad_2 calculation strategy.

In addition, we introduce a dual-population strategy, which divides the ACO into two populations. The first population focuses on the global search, and the second population focuses on local search. The calculation formula is defined as follows:

$$\theta_{(i,x,y)} = \begin{cases} \exp(-|r_{pre} + x - 2r| - |s_{pre} + y - 2s|), & i = 1 \\ \exp(|r_{pre} + x - 2r| + |s_{pre} + y - 2s|), & i = 2 \end{cases}, \quad (6)$$

where $\theta_{(i,x,y)}$ is a directional factor controlling the movement direction of the ant of the population i and γ is an influence factor of the azimuth factor $\theta_{(i,x,y)}$. i ($i=1, 2$) represents the population number, (r, s) is the current position of the ant, (r_{pre}, s_{pre}) is the previous position of the ant, and (x, y) is the next position of the ant.

The double-population strategy ensures that the ants of the first population will advance in the same direction as the original direction with a high probability and perform a global search to find the optimal solution as soon as possible in the initial stage of the search. The ants of the second population will advance in the same direction as the original direction with a small probability, move around the current position with a high probability, and perform a local search to carefully find the optimal solution near the optimal solution.

4.3 Pheromone Update and Edge Extraction

4.3.1 Pheromone update

After the ant moves one step each time, the artificial ant updates the pheromone on the pixel it visits using the following Eq. (7).

$$\tau_{(x,y)}^n = (1 - \rho)\tau_{(x,y)}^{n-1} + \rho \sum_k \eta_{(x,y,k)}, \quad (7)$$

where ρ is the evaporation coefficient. The reward after an ant moves to a pixel (x, y) is defined as the difference between the pheromone of the pixel (x, y) and its neighboring pixel:

$$\text{reward} = \sum_{u,v=-1,0,1} (\tau_{(i,j)} - \tau_{(i+u,j+v)}). \quad (8)$$

We update the Q function using Eq. (1). After the ant moves one step, the pheromone matrix is updated by Eq. (9):

$$\tau^n = (1 - \varphi)\tau^{n-1} + \varphi\tau^0, \quad (9)$$

where φ represents the attenuation coefficient of the pheromone and τ^0 is the initial pheromone matrix. φ is used to avoid the ants in the path from depositing pheromone to avoid local optimization. When the threshold number of iterations is reached, the search is terminated, and the final pheromone matrix can then be used to detect edges in the image.

4.3.2 Edge extraction

The intensity threshold based on the pheromone matrix

is calculated as follows:

Step 1: Select the average value T in the pheromone matrix and use T as the initial value of the threshold.

Step 2: Divide the pheromone matrix into two groups: the first group contains elements larger than T , and the other group contains elements smaller than T .

Step 3: Calculate the average of the two groups, namely, T_1 and T_2 , and let $T = (T_1 + T_2)/2$.

Step 4: Repeat the above three steps until T is not changed, and then obtain T as the threshold. If the value of the element in the pheromone matrix is greater than T , the corresponding pixel belongs to the edge; otherwise, the pixel does not belong to the edge.

5 Experiments

5.1 Experimental Test Environment

During the experiment, the computer hardware environment is as follows: Intel Core i9-9900k processor, GeForce RTX 2080 Ti graphics card, 64GB memory, Windows 10 (64bit) system type, and Matlab R2018b program development tool. The pheromone influence factor α is 6, the heuristic information influence factor β is 0.1, the orientation influence factor γ is 1, the initial pheromone τ_0 is 0.0001, and the maximum number of iterations is 1000.

5.2 Experimental Test Plan

First, two typical methods are selected to compare the methods proposed in this paper. The first ACO-based method uses the fixed perception radius 1 to calculate the gradient in heuristic information, instead of using the direction factor, herein referred to as ‘perception radius 1’ in short (Nezamabadi-pour *et al.*, 2006). The second ACO-based method uses the fixed sensing radius 2 to calculate the gradient fixed sensing radius 2 in heuristic information, herein referred to as ‘sensing radius 2’ in short (Tian *et al.*, 2008). For each test image, the edge detection results of each method are given.

The real image taken by an underwater camera (DEEP-SEA Power & Light, San Diego, USA) in the turbid medium is used as the edge detection dataset. Three images are randomly selected: underwater badminton (Fig.5), fish (Figs.6 and 7), and turtle (Figs.8 and 9). Figs.5–9 show the image edge detection results. In each figure, the subgraph (a) is the original image, (b) shows the image edge detection results of the ‘perception radius 1’ strategy, (c) shows the image edge detection results of the ‘perception radius 2’ strategy, and (d) shows the image edge detection results of the proposed method.

Table 1 shows the comparison values of the quantitative evaluation indexes of each method on the test image. Quantitative criteria are used to evaluate the methods proposed in this paper. The quantitative criteria include completeness, discriminability, precision, and robustness (Moreno *et al.*, 2009). The completeness index measures the ability of edge detection to detect all possible edges of a noiseless image. Discriminability measures the ability of edge detection to distinguish an important edge from an unim-

portant edge. Precision measures the degree to which the edge detected by the edge detection method is close to the

actual edge. Robustness measures the ability of edge detection to be immune to noise.

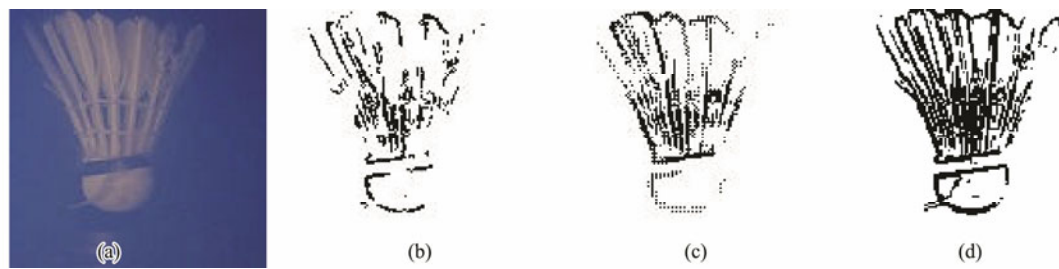


Fig.5 Edge detection results for the badminton image. (a), real image; (b), perception radius 1; (c), perception radius 2; (d), proposed method.

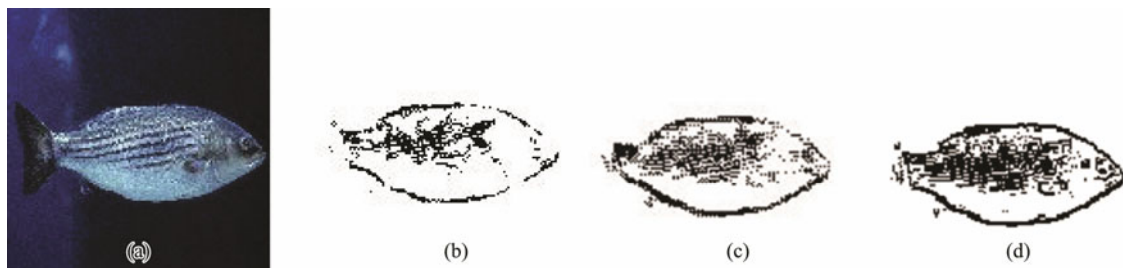


Fig.6 Same as Fig.5 but for the fish image.

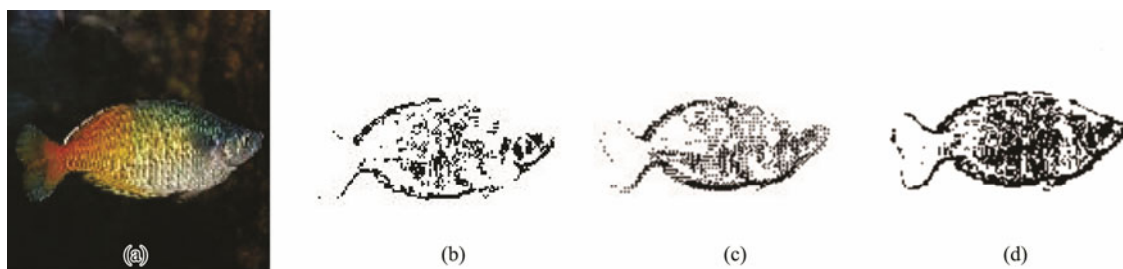


Fig.7 Same as Fig.5 but for the fish image.

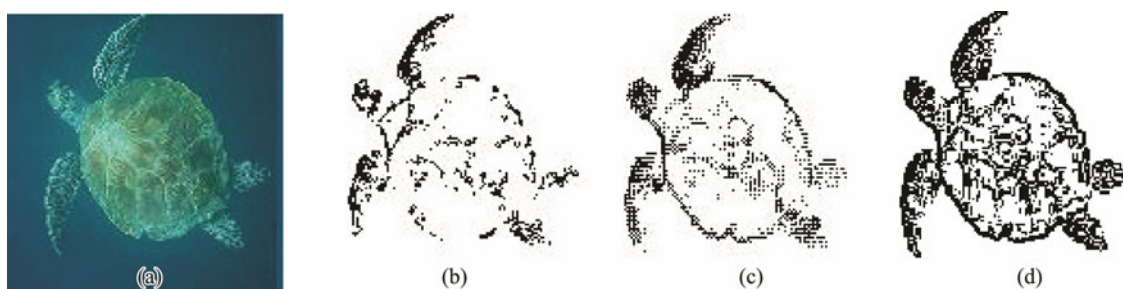


Fig.8 Same as Fig.5 but for the turtle image.

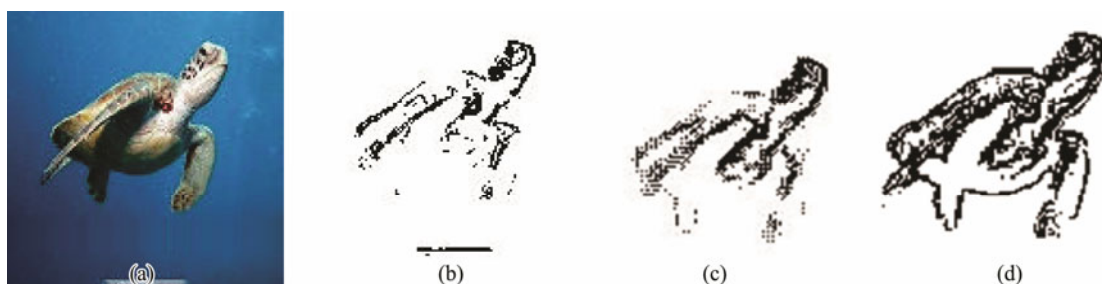


Fig.9 Same as Fig.5 but for the turtle image.

Table 1 Quantitative comparison results of the methods

	Perception radius 1	Perception radius 2	Proposed method
Completeness	0.70	0.78	0.93
Discriminability	11.60	13.21	16.81
Precision	0.70	0.79	0.85
Robustness	0.58	0.51	0.63

As shown in Figs.5 to 9 and Table 1, the method proposed in this paper can detect the unclear edge pixels and can also remove some noise in the image, which is superior to the comparison method in details and as a whole.

Fig.10 shows the pheromones deposited on the pixels of the image after the end of the algorithm. The plane in the two horizontal axes represents the image, and the vertical axis represents the pheromone on each pixel. Fig.10(a) shows the pheromone of Fig.7, and Fig.10(b) shows the pheromone of Fig.8. The experimental results show that the algorithm can extract the image edge effectively and deposit the pheromone on the edge pixels of the image correctly.

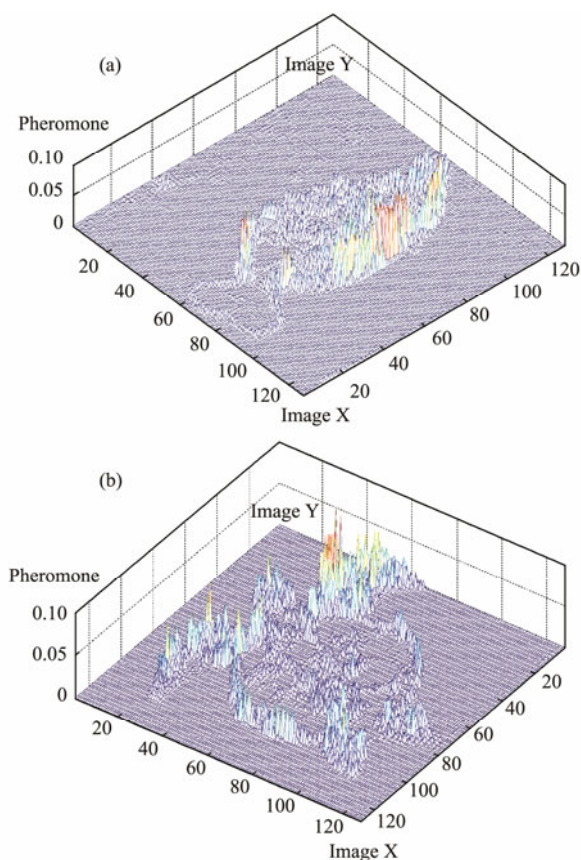


Fig.10 Pheromone of ants deposited on images. (a), the pheromone is deposited on the image in Fig.7. (b), the pheromone is deposited on the image in Fig.8.

Finally, to evaluate the performance of the algorithm objectively and accurately, 20 images are randomly selected from the actual underwater image dataset as test samples and then randomly divided into 5 groups with 4 images in each group. The ‘perceptual radius 1’ method, ‘perceptual radius 2’ method, and the algorithm proposed

in this paper are used to evaluate the image edge detection algorithm quantitatively. The comparison results are shown in Tables 2–5.

Table 2 Completeness of the comparison results of the methods

Sample group	Perception radius 1	Perception radius 2	Proposed method
1	0.73	0.75	0.81
2	0.65	0.68	0.79
3	0.81	0.82	0.90
4	0.88	0.83	0.95
5	0.79	0.79	0.86

Table 3 Discriminability comparison results of the methods

Sample group	Perception radius 1	Perception radius 2	Proposed method
1	10.80	12.31	15.63
2	8.05	9.56	13.91
3	12.37	12.15	16.94
4	18.65	17.56	19.89
5	12.54	11.32	16.23

Table 4 Precision comparison results of the methods

Sample group	Perception radius 1	Perception radius 2	Proposed method
1	0.72	0.81	0.86
2	0.56	0.66	0.76
3	0.82	0.81	0.82
4	0.91	0.93	0.94
5	0.80	0.78	0.78

Table 5 Robustness comparison results of the methods

Sample group	Perception radius 1	Perception radius 2	Proposed method
1	0.59	0.61	0.65
2	0.43	0.43	0.45
3	0.67	0.71	0.75
4	0.70	0.71	0.74
5	0.64	0.62	0.67

6 Conclusions

In this paper, an ant colony pheromone computing strategy based on reinforcement learning is proposed, which is different from the traditional gradient computing method with a fixed number of neighborhood pixels and can realize adaptive variable neighborhood pixel gradient computing. In addition, a double-population strategy is proposed to control the movement direction of ants: the first population focuses on the global search, whereas the second population focuses on the local search. Thus, the search process takes into account global search and local search abilities. Experimental results show that the algorithm can effectively reduce the impact of noise on edge detection, obtain complete and clear image edges, and achieve good results. In the future, we will study the complexity of the algorithm to further improve its efficiency and adaptively adjust other parameters in the algorithm.

Acknowledgements

The study is supported by the start-up fund for doctoral research of Northeast Electric Power University (No. BS JXM-2020219) and the Science and Technology Research Program of the Jilin Provincial Department of Education (No. JJKH20210115KJ).

References

- Bonin-Font, F., Oliver, G., Wirth, S., Massot, M., Negre, P., and Beltran, J., 2015. Visual sensing for autonomous underwater exploration and intervention tasks. *Ocean Engineering*, **93**: 25-44.
- Chuang, M., Hwang, J., and Williams, K., 2016. A feature learning and object recognition framework for underwater fish images. *IEEE Transactions on Image Processing*, **25** (4): 1862-1872.
- Dawson, L., and Stewart, I., 2014. Accelerating ant colony optimization-based edge detection on the GPU using CUDA. *2014 IEEE Congress on Evolutionary Computation*. Beijing, 1736-1743.
- Deng, W., Xu, J. J., and Zhao, H. M., 2019. An improved ant colony optimization algorithm based on hybrid strategies for scheduling problem. *IEEE Access*, **7**: 20281-20292.
- Fatan, M., Daliri, M., and Shahri, A., 2016. Underwater cable detection in the images using edge classification based on texture information. *Measurement*, **91**: 309-317.
- Gao, Y. Q., Guan, H. B., Qi, Z. W., Hou, Y., and Liu, L., 2013. A multi-objective ant colony system algorithm for virtual machine placement in cloud computing. *Journal of Computer and System Sciences*, **79** (8): 1230-1242.
- Kheirinejad, S., Hasheminejad, S., and Riahi, N., 2018. Max-min ant colony optimization method for edge detection exploiting a new heuristic information function. *International Conference on Computer and Knowledge Engineering*. Mashhad, 12-15.
- Kober, J., Bagnell, J., and Peters, J., 2013. Reinforcement learning in robotics: A survey. *International Journal of Robotics Research*, **32** (11): 1238-1274.
- Lewis, F., and Vrabie, D., 2009. Reinforcement learning and adaptive dynamic programming for feedback control. *IEEE Circuits and Systems Magazine*, **9** (3): 32-50.
- Li, Q., Sun, X., Dong, J. Y., Song, S. Q., Zhang, T. T., Liu, D., *et al.*, 2020. Developing a microscopic image dataset in support of intelligent phytoplankton detection using deep learning. *ICES Journal of Marine Science*, **77** (4): 1427-1439.
- Lin, Y. H., Chen, S. Y., and Tsou, C. H., 2019. Development of an image processing module for autonomous underwater vehicles through integration of visual recognition with stereoscopic image reconstruction. *Marine Science and Engineering*, **7** (4): 107-149.
- Liu, F., Wei, Y., Han, P. L., Yang, K., Lu, B., and Shao, X. P., 2019. Polarization-based exploration for clear underwater vision in natural illumination. *Optics Express*, **27** (3): 3629-3641.
- Liu, X. C., and Fang, S. P., 2015. A convenient and robust edge detection method based on ant colony optimization. *Optics Communications*, **353**: 147-157.
- Lu, D. S., and Chen, C. C., 2008. Edge detection improvement by ant colony optimization. *Pattern Recognition Letters*, **29** (4): 416-425.
- Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., *et al.*, 2015. Human-level control through deep reinforcement learning. *Nature*, **518** (7540): 529-533.
- Montague, P., 1999. Reinforcement learning: An introduction, by Sutton, R. S., and Barto, A. G. *Trends in Cognitive Sciences*, **3** (9): 360-360.
- Moreno, R., Puig, D., Julia, C., and Garcia, M., 2009. A new methodology for evaluation of edge detectors. *IEEE International Conference on Image Processing*. Cairo, 2157-2160.
- Nasution, T., Zarlis, M., and Nasution, M., 2017. Optimizing robinson operator with ant colony optimization as a digital image edge detection method. *International Conference on Information and Communication Technology*. Medan, 930012034.
- Nezamabadi-pour, H., Saryazdi, S., and Rashedi, E., 2006. Edge detection using ant algorithms. *Soft Computing*, **10** (7): 623-628.
- Rafsanjani, M. K., and Varzaneh, Z. A., 2015. Edge detection in digital images using ant colony optimization. *Computer Science Journal of Moldova*, **23** (3): 343-359.
- Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., *et al.*, 2018. A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science*, **362** (6419): 1140-1144.
- Singh, R., Vashishath, M., and Kumar, S., 2019. Ant colony optimization technique for edge detection using fuzzy triangular membership function. *International Journal of Systems Assurance Engineering and Management*, **10** (1): 91-96.
- Sun, X., Shi, J. Y., Liu, L. P., Dong, J. Y., Plant, C., Wang, X. H., *et al.*, 2018. Transferring deep knowledge for object recognition in low-quality underwater videos. *Neurocomputing*, **275**: 897-908.
- Tian, J., Yu, W. Y., and Me, S. L., 2008. An ant colony optimization algorithm for image edge detection. *IEEE Congress on Evolutionary Computation*. Hong Kong, 751-756.
- Vinyals, O., Babuschkin, I., Czarnecki, W. M., Mathieu, M., Dudzik, A., Chung, J., *et al.*, 2019. Grandmaster level in StarCraft II using multi-agent reinforcement learning. *Nature*, **575** (7782): 350-370.
- Wang, X. H., Ouyang, J. H., Li, D. Y., and Zhang, G., 2019. Underwater object recognition based on deep encoding-decoding network. *Journal of Ocean University of China*, **18** (2): 376-382.

(Edited by Chen Wenwen)