Autonomous Path Planning Scheme Research for Mobile Robot

Jianxian Cai[1,2], *Xiaogang Ruan*[1,3]*, Pengxuan Li*[2]

[1]*School of Electronic and Control Engineering, Beijing University of Technology, No 100, Pingleyuan, Chaoyang District, 100124 Beijing, China*
[2]*Department of Disaster Prevention Instrument, Institute of Disaster Prevention, No 3, University Street, Yanjiao Development Zone, Sanhe, 065201 Hebei, China*
[3]*Beijing Key Laboratory of Computational Intelligence and Intelligent System, No 100, Pingleyuan, Chaoyang District, 100124 Beijing, China*
Emails: cjxlaq@163.com

Abstract: *An autonomous path-planning strategy based on Skinner operant conditioning principle and reinforcement learning principle is developed in this paper. The core strategies are the use of tendency cell and cognitive learning cell, which simulate bionic orientation and asymptotic learning ability. Cognitive learning cell is designed on the base of Boltzmann machine and improved Q-Learning algorithm, which executes operant action learning function to approximate the operative part of robot system. The tendency cell adjusts network weights by the use of information entropy to evaluate the function of operate action. The results of the simulation experiment in mobile robot showed that the designed autonomous path-planning strategy lets the robot realize autonomous navigation path planning. The robot learns to select autonomously according to the bionic orientate action and have fast convergence rate and higher adaptability.*

Keywords: *Autonomous path planning, cognitive learning, information entropy, Q-Learning, mobile robots.*

1. Introduction

Mobile robot is a kind of a robot system that can perceive its environment and its own state through sensors. Sensors give the robot capability of moving autonomously targeting an environment with obstacles for accomplishing a certain task. In recent years, the technology of autonomous navigation of mobile robot becomes ever demanding with the development of science and technology. One of the demands is the autonomous path planning, which is considered as a core technology, and at the same time, a challenge in the mobile robot autonomous navigation technology. This means that robots can autonomously learn an optimal or suboptimal path from initial state to goal state according to a certain performance index [1].

At present, autonomous path planning in unknown environments gained less research. Main methods are artificial potential field method, neural networks method, genetic algorithm method and fuzzy logic method. Artificial potential field method [2] has simple structure and is easy to be implemented. However, there are pitfalls that this method experiences, such as its inability to find the path in front of similar obstacles, being prone to oscillation in front of a barrier and its tendency to swing in a narrow passage. Neural networks method [3-7] generally adopts supervised training to optimize parameters, and is susceptible to some constraints such as difficulty in obtaining learning supervision signal and the likes. Genetic algorithm method's [8] search strategy and optimization calculation is not dependent on the gradient information, thus, it can solve some problems that other optimization algorithms cannot. Nevertheless, there are still disadvantages in long calculation time and it also experiences difficulty in online optimization. Fuzzy logic method [9-11] expresses knowledge with fewer rules by making full use of field knowledge, and so shows great superiority in dealing programming problem in unknown environments. However, rule explosion will only occur when the environment is more complicated [12].

With the research getting further especially and with constantly expanding applications in the field of mobile robot, the emphasis are on robot intelligence level of perception; strain on the environment is gaining much attention. As a result, extensive research of autonomous mobile robot path planning is carried out by simulating cognition, behavior and learning biological mode. W a n g and C e n g [13] of China University of Geosciences applied reinforcement learning to inverted pendulum control; experimental verification and analysis were carried out. A method of mobile robot reactive navigation based on reinforcement learning is put forward by X u-X i [14] of National University of Defense Technology; it was successfully applied to the actual CIT-AVT-VI mobile robot platform and overcome the local minimum compared with the artificial potential field method. Y e, Y u n g and W a n g [15] designed a kind of fuzzy controller combined with supervision learning and reinforcement learning, which was successfully applied on mobile robot obstacle avoidance. But the controller has 243 rules and so it is difficult to satisfy the real-time requirement. B o u b e r t a k h, T a d j i n e and G l o r e n n e c [16] put forward an improved fuzzy Q-Learning method, which has only eight fuzzy rules based on expert knowledge and has successfully completed the mobile robot obstacle avoidance and target navigation task. However, the method can only implement the perception-match action learning and it lacks the ability to solve complex problems. In addition, some scholars have studied other bionic learning model. For example, classical conditioned reflex and operation conditioning was blend in behavior learning of robot by C h a n g and G a u d i a n o [17] of Boston University Neurobotics Laboratory; robot can learn phototaxis and obstacle avoidance behavior in the process of navigation without any prior knowledge and training.

In conclusion, some research achievements have been made on the study of the cognitive ability of the autonomous mobile robot path planning. Among them, the reinforcement learning algorithm has highlighted the advantages; it provides

referent frame for robot autonomous path plan, however, it cannot reflect the reality structure and function of biological nervous system, so there are still many problems yet to be solved. Operating conditions have achieved ideal results in the application of the robot autonomous path planning, but most of the cognitive computing models are descriptive, abnormal and do not have generalization ability. Therefore, a calculated cognitive model will be established by simulating biology operant conditioning in the framework of mature reinforcement learning for implementing robot autonomous path planning.

2. Autonomous path planning model design

Fig. 1 shows the mobile robot autonomous navigation diagram by simulating operant conditioning mechanism. State editor module solves the problem of dividing environment states; navigation learning strategy module solves the problem of mapping from state space to action space; environment interaction module solves the problem of producing evaluated signal and state transfer. Mobile robot keeps interacting with the environment in the process of navigation learning and anticipative object will be achieved through learning and training.
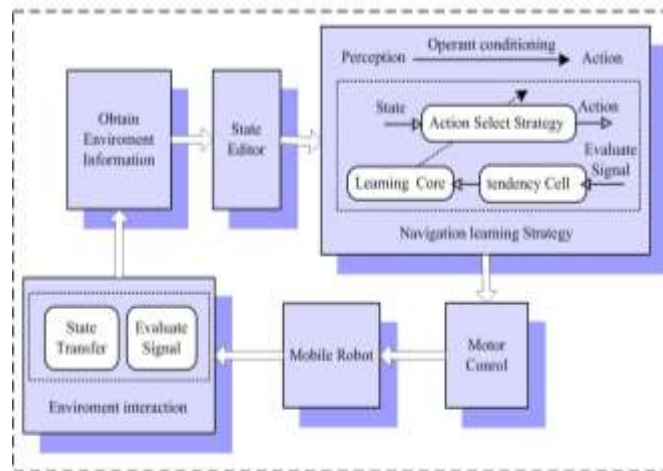

Fig. 1. Navigation strategy system of simulating operant conditioning

2.1. Division of environment state

Robot continuously learns in the process of navigation, and environment state will be transferred after learning. Learning produces new interaction, and the new interaction will stimulate further learning. Robot action will become better and better and will gradually adapt the environment through reiteration. System state should be decomposed and should form state space for promoting environment interaction.

State and action of mobile robot are considered in local unknown environment, and the navigation aim is the robot to reach the goal point from start point free of collision. The relationship schematic diagram between a robot and its environment

is shown in Fig. 2. The robot is equipped with sixteen sonar sensors, and the distance among sonar sensor is 20° or 40°. The coverage area of sonar sensors contains 0-360° and so, robot can freely rotate in the area of 0-360°.
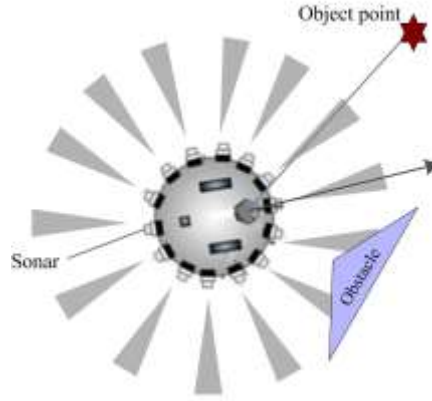

Fig. 2. Relationship schematic diagram between robot and environment

To conveniently study navigation arithmetic, robot is supposed to continuously move with equal space and each step size is a grid length. At the same time, robots can freely turn round without collision in narrow environment area. Hence, robot is simplified as a particle and does not consider the rotation radius in navigation arithmetic. The relation between the robot and the obstacle object is shown in Fig. 3.
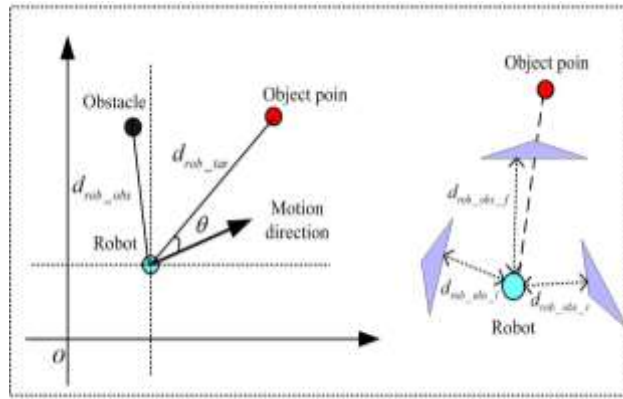

Fig. 3. Position relationships among the robot, obstacle and goal

The modest way of sensor information fusion is used in order to simplify the problem and reduce the size of the state space. The detection range of the sensor is divided into three areas (left, frontage and right) to reduce state space. Having this, the environmental state between the robot and the obstacles is shown by three distances information of direction. In addition, considering the relation between the robot and the object, the distance information between the robot and the target point, and an angle between the targets are added into state space.

Definition 1. Robot state space s is

(1) $$s = \{d_{r\_l}, d_{r\_f}, d_{r\_r}, d_{r\_tar}, \theta\},$$

where d_{r_l} is the distance between the robot and left obstacles;

d_{r_f} is the distance between the robot and front obstacles;

d_{r_r} is the distance between the robot and right obstacles;

d_{r_tar} is the distance between the robot and the target;

θ is the angle between the robot movement direction and the target.

The distance of three directions discrete to N (Near) and F (Far) are

(2) $$S = \{d_{r\_l}, d_{r\_f}, d_{r\_r}\} = \begin{cases} N, & d_s \leq \min(d_{r\_l}, d_{r\_f}, d_{r\_r}) < d_m, \\ F, & \min(d_{r\_l}, d_{r\_f}, d_{r\_r}) > d_s, \end{cases}$$

where d_s is the smallest dangerous distance, and the robot will fail to avoid obstacle when any detection range is less than d_s; d_m is the largest safe distance and the robot will walk safely with maximum speed when any detection range is greater than d_s.

2.2. Action space

The control variables of robot are linear velocity v and kinematic angle $\Delta\theta$. Robot can realize obstacle avoidance and goal orientation if v and $\Delta\theta$ are appropriate, and so, the action space of robot is defined as:

(3) $$a = \{v, \Delta\theta\},$$

where v and $\Delta\theta$ are describe as: PB (Positive Big), PM (Positive Medium), PS (Positive Small), ZR (Zero), NS (Negative Small), NM (Negative Medium) and NB (Negative Big).

2.3. Learning rules

Learning rules are built on the base of expert knowledge about the task of obstacle avoidance and goal orientation, which can greatly reduce the search space of the robot so it can respond quickly to its environment.

Expert knowledge is described as:

(1) Robot will navigate towards the target with the greatest speed if it stays away from obstacles in three directions;

(2) Robot moves left or right towards the target's direction along the nearest obstruction if the target is not in front of the robot, but the obstacles is in front;

(3) Robot tries to move left or right towards the target's direction along the nearest obstruction if the target and obstacles are in front of the robot.

Two parameters, ρ and q, are set so that the robot can decide autonomously the path planning model:

(4) $$\rho = \begin{cases} 0 & \text{if } d_{r\_l} \leq d_{r\_r} \text{ and } d_{r\_l} < d_m, \\ 1 & \text{else}, \end{cases}$$

(5) $$q = \begin{cases} 0 & \text{if } 90° < \theta < 270°, \\ 1 & \text{else}, \end{cases}$$

117

where the detectable maximum distance is d_{m} which is assumed as 5 m. The robot walks right (left) along the left wall when $\rho = 1$, $\rho = 0$, activates the goal orientation behavior when $q=1$, and walks along the wall when $q=0$.

Eight learning rules are produced according to the definition of state space and action space:

R1: IF $(d_{\mathrm{r\_l}}, d_{\mathrm{r\_f}}, d_{\mathrm{r\_r}})$ is (N, N, N) THEN v is ZR and $\Delta\theta$ is $(2\times\rho-1)\times$PB;

R2: IF $(d_{\mathrm{r\_l}}, d_{\mathrm{r\_f}}, d_{\mathrm{r\_r}})$ is (N, N, F) THEN v is ZR and $\Delta\theta$ is NB;

R3: IF $(d_{\mathrm{r\_l}}, d_{\mathrm{r\_f}}, d_{\mathrm{r\_r}})$ is (N, F, N) THEN v is $v_{\max}$ and $\Delta\theta$ is ZR;

R4: IF $(d_{\mathrm{r\_l}}, d_{\mathrm{r\_f}}, d_{\mathrm{r\_r}})$ is (N, F, F) THEN v is $V_{\max}$ and $\Delta\theta$ is $(2\times\rho-1)\times$NS;

R5: IF $(d_{\mathrm{r\_l}}, d_{\mathrm{r\_f}}, d_{\mathrm{r\_r}})$ is (F, N, N) THEN v is ZR and $\Delta\theta$ is PB;

R6: IF $(d_{\mathrm{r\_l}}, d_{\mathrm{r\_f}}, d_{\mathrm{r\_r}})$ is (F, N, F) THEN v is ZR and $\Delta\theta$ is $(2\times\rho-1)\times$NB;

R7: IF $(d_{\mathrm{r\_l}}, d_{\mathrm{r\_f}}, d_{\mathrm{r\_r}})$ is (F, F, N) THEN v is $V_{\max}$ and $\Delta\theta$ is $(2\times\rho-1)\times$NS;

R8: IF $(d_{\mathrm{r\_l}}, d_{\mathrm{r\_f}}, d_{\mathrm{r\_r}})$ is (F, F, F) THEN v is $V_{\max}$ and $\Delta\theta$ is $q\times\theta+(1-q)\times(2\times\rho-1)\times$PM;

here $V_{\max}$ is the maximum speed. The moving speed v and kinematic angle $\Delta\theta$ of the robot corresponding to each state can be obtained by above eight rules, the value of each variable PB, PM, PS, ZR, NS, NM, NB are adjusted by using each learning strategy.

3. Realization of autonomous path planning

3.1. Realization of autonomous learning strategy

(1) Bionic autonomous learning strategy

Complexity of bionic autonomous learning strategy will index increase with the increase of states number. The related literature pointed out that states related to the best effect are lower than a quarter. Therefore, information entropy is introduced, which can weigh the state tendency by computing entropy value of environment state. Then, the information entropy of state s_i is defined as

$$(6) \qquad H_i(t) = -\sum_{k=1}^{r} p_{ik}\log_2 p_{ik},$$

where $a_k \in A = \{a_k \mid k = 1, 2, \cdots, r\}$ is the k-th selected action; A is the selected action set; r is the number of selected action; $p_{ik} = p(a_k \mid s_i) \in P_i$ represents the probabilistic value of operant action a_k with state s_i and also is called excited probability value of "state s_i-action a_k" pare, which satisfies $0 < p_{ik} < 1$, $\sum_{k=1}^{r} p_{ik} = 1$;

$P_i = \{p_{i1}, p_{i2}, \ldots, p_{ir}\} \in P$, P_i represents the probability vector of the i-th state and P represents the overall probability vector.

118

(2) Designation of the learning core

The learning core includes designation of evaluation signal and reinforcement function. An evaluation signal will be generated after any action to evaluate the effect of a selected action. The reasonable distribution of the reward and punishment evaluation signal is helpful in improving the learning efficiency. In addition, considering mobile robot path planning in unknown environment exploration has two goals, the obstacle avoidance and the approaching goal, so the design of evaluation signal should consider several aspects.

1） $d_{r\_obs} > d_m$.

Parameter d_{r_obs} indicates the distance between the robot and the obstacles. There is a little possibility of collision between the robot and the obstacles when d_{r_obs} is greater than the safe distance and the main task of the robot is to reach its target point. Consequently, the key design of evaluation signal will mainly consider the relation between the distance d_{r_tar} and the angle:

robot turns to the goal direction, then $\begin{cases} \Delta\theta(t) = \theta(t+1) - \theta(t), \\ \qquad \Delta\theta(t) < 0; \end{cases}$

robot is expected to approach the goal, then $\begin{cases} \Delta d_{r\_tar}(t) = d_{r\_tar}(t+1) - d_{r\_tar}(t), \\ \qquad \Delta d_{r\_tar}(t) < 0. \end{cases}$

Then the evaluated signal can be written as

(7) $\qquad V(d_{r\_tar}, \theta) = -\beta_1 \text{sign}\left(\Delta d_{r\_tar}(t)\right)\Delta^2 d_{r\_tar}(t) - \beta_2 \text{sign}\left(\Delta\theta(t)\right)\Delta^2\theta(t),$

where β_1 and β_2 are weight coefficients, and $0 < \beta_1, \beta_2 < 1$.

2） $d_s < d_{r\_obs} < d_m$.

Robot will possibly collide the obstacles when d_{r_obs} is between dangerous distance and safe distance, and the main task of the robot is to reach the target point and avoid the obstacle. Therefore, the key design of evaluation signal will mainly consider the relation between distance d_{r_tar} and d_{r_obs},

Robot is expected to avoid obstacle, then $\begin{cases} \Delta d_{r\_obs}(t) = d_{r\_obs}(t+1) - d_{r\_obs}(t), \\ \qquad d_{r\_obs}(t) > 0. \end{cases}$

The evaluation signal can be written as

(8) $V(d_{r\_obs}, d_{r\_tar}) = \beta_3 \text{sign}\left(\Delta d_{r\_obs}(t)\right)\Delta^2 d_{r\_obs}(t) - \beta_4 \text{sign}\left(\Delta d_{r\_tar}(t)\right)\Delta^2 d_{r\_tar}(t),$

where β_3 and β_4 are weight coefficients, and $0 < \beta_3, \beta_4 < 1$.

3） $d_{r\_obs} < d_s$.

The biggest value of the punishment value will be set to $V = -1$ when distance between the robot and the obstacle is smaller than dangerous distance.

Q-function is designed on the base of evaluation signal as follows:

(9) $Q(s_i(t), a_k(t)) = \left(1 - \gamma(p_{ik})\right)Q(s_i(t-1), a_k(t-1)) + \gamma(V_t)\left[V_t + \eta \max_{a_k} Q(s_i(t+1), a_k(t))\right],$

η is the discounted factor, which represents attentive degree of learning system to action. Learning system concerns the lately action when the value of discounted factor is small. Learning rate function $\gamma(p_{ik})$ controls the learning velocity of system. The value $\gamma(p_{ik})$ is bigger, while the convergence rate is faster. However, excessive value of $\gamma(p_{ik})$ will cause non-convergence. Thus, the learning rate function can be designed as

(10)
$$\gamma(p_{ik}) = \frac{1}{1 + \exp\left[\dfrac{p_{ik}(t) - p_{ik}(t+1)}{p_{ik}(t)}\right]}.$$

(3) Designation of action selected strategy

Action selected strategy is the core part of robot navigation. The main task of the robot is to explore its environment in the early period of learning, so randomness of selecting action is big; the main task of robot is ensuring learning convergence in the later period of learning, so the randomness of selecting action is small. Boltzmann machine is adopted to realize annealing calculation. The selected probability of action a_k is

(11)
$$p_{ik}(s, a) = \frac{e^{Q(s,a)/T}}{\sum_{a_k \in A} e^{Q(s,a_t)/T}}, \quad T = T_0 t^{-1/\varphi},$$

where T is the temperature coefficient, T_0 is the initial temperature value, parameter φ is used to control the velocity of anneal.

3.2. The process of learning algorithm

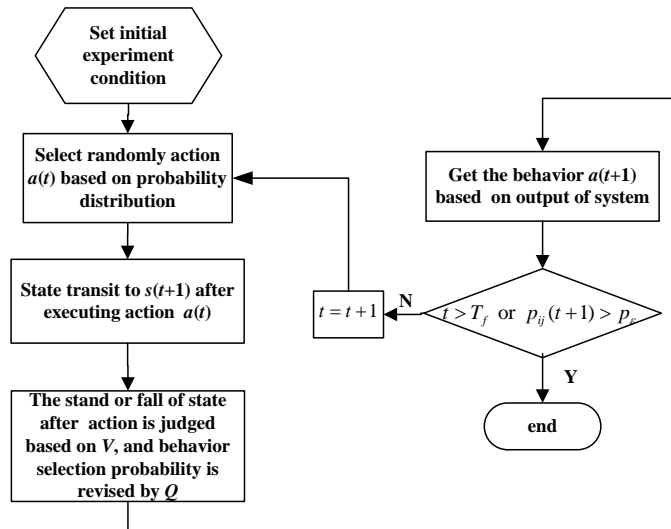The process of learning algorithm is shown in Fig. 4.

Fig. 4. The autonomous learning algorithm flowchart

4. Experimental results and analysis

4.1. Experimental conditions

In the platform MobotSim, mobile robot working environment was created for validating the feasibility of designed autonomous path-planning strategy. The environment is shown by a grid map consisting of 70×50 grids, and the size of each grid is 0.2×0.2 m. There is a static obstacle in the navigation environment. Mobile robot can detect different direction of obstacle distance by sonar sensors, perceived the current location and the target location information, and then found the optimal or the best path to arrive at the target point by autonomous learning. Robot is represented by a red little roundness with diameter about 0.5 m, and the goal, which is set to be yellow, is also a little roundness.

The purpose of the paper is to make the mobile robot to be capable of obstacle avoidance and target orientation by learning like person or animal, and at the same time the mobile robot is expected to overcome the local minimum and oscillation problem.

The initial iterative learning step is t=0; the weight coefficients are $\beta_1 = 0.65$, $\beta_2 = 0.36$, $\beta_3 = 0.72$ and $\beta_4 = 0.44$; the discounted factor is $\eta = 0.14$; the minimum dangerous distance is d_s=0.1 m, the biggest safety distance is d_m=0.9 m; the maximum mobbing speed of the robot is $V_{\max}$ =0.15 m/s; the action space is $A = \{a_k \mid k = 1, 2, \cdots, 7\}$; the selected probability of each action is about $p_{ik}(0) \approx \dfrac{1}{7}$ and correspondingly the initial entropy value of each state is maximum with value

$$H_i(0) = -\sum_{k=1}^{5} p_{ik} \times \log_2 p_{ik} \bigg|_{p_{ik} = \frac{1}{5}} \approx 2.72 .$$

Table 1. The symbol value sets of $\Delta\theta$

Symbol value	NS	NM	NB	ZR	PB
MIN	−20	−40	−60	0	40
MAX	−5	−20	−40	0	60

4.2. Results and analysis

Experiment 1. Target path planning experiments

The autonomous path planning experiment is preceded in obstacles sparse environment. A collision will be recorded when the distance between the robot and the obstacles is less than the minimum dangerous distance; after that, the robot will select an action again until it becomes aware of environment and could avoid obstacles. Tendency cell will evaluate the executed action, it is used to the cognitive learning unit, and then the probability of better action will be updated.

Training result of the initial and ending training phase was taken out for clearly showing the bionic learning process of self-learning and self-organization (see Fig. 7). There are eleven irregular shape obstacles in simulation environment

and the yellow sector part is the coverage area of sensor. The starting point is (10, 10) and the target point is (55, 35).

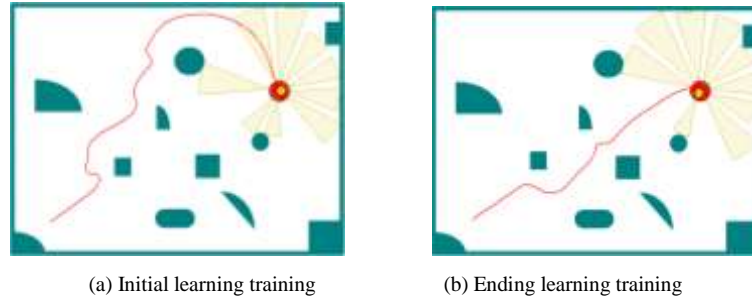(a) Initial learning training   (b) Ending learning training

Fig. 5. Path planning trajectory based on bionic autonomous learning strategy

The simulation result of Fig. 5a showed that the path planning trajectory wasn't at its optimal; collision occurs once during the initial learning because the robot hadn't any prior knowledge but had high random performance. So, the robot could not select effective action to execute path planning. With the interaction between the robot and its environment, the robot began incline to select the action which could obtain more reward and improve the performance of path planning. The simulation result of Fig. 5b showed that the robot could move to the target point, it avoided static obstacle in the same time of terminal learning, which showed the effective action, so the collision free path planning had been learned.

The simulation result of Figs 6 and 7 are time steps and collision frequency curve of 20 times test respectively. The environment in the initial learning is unknown for the robot, so it can be seen in the simulation curves that the robot makes collision 7 times in first learning and reaches the target after 300 time steps. The collision frequency has been reduced to 0 and only need 115 time steps after 3 times learning. The robot can reach the target without collision and use only 48 time steps after 20 times learning. The simulation results show that the robot can learn autonomously in obstacles sparse environment and can complete the autonomous path planning task.
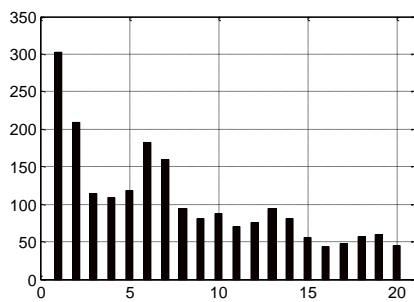

Fig. 6. Curves of the relationship between the number of learning and time steps for navigation
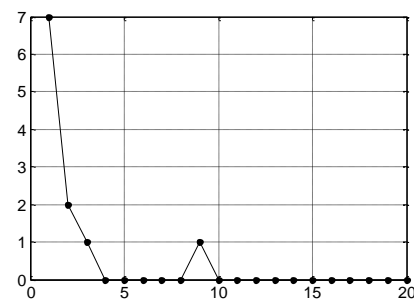
Fig. 7. Curves of he relationship between the number of learning and collision for navigation

The learned "state-action" mapping relations are saved and the working environment is changed by setting different starting point and target point. The

distributions of obstacles are also changed, and the path planning experiment was conducted again. The simulation result could be seen on Fig. 8, which shows that the robot could evade obstacles successfully; it arrived at goal point in the different environment by using designed bionic learning strategy. The learning results established the validity of the designed autonomous path planning strategy.
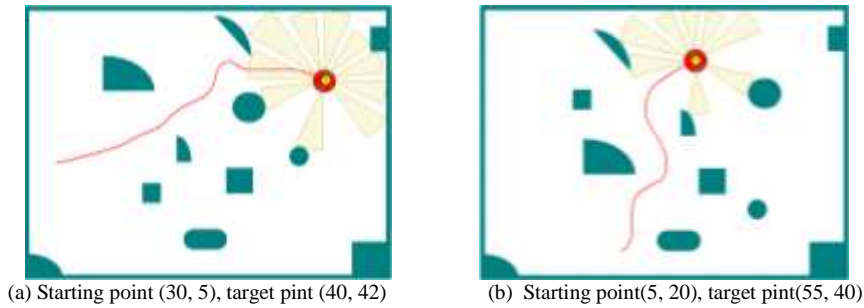

(a) Starting point (30, 5), target pint (40, 42)   (b) Starting point(5, 20), target pint(55, 40)

Fig. 8. Navigation trajectory after change the environment

Experiment 2. Contrast experiment of overcoming local minimum and oscillation around

In order to further verify the validity of autonomous path planning strategy, bionic leaning strategy and artificial potential field method [10] were compared in the same environment that contains double U obstacles and narrow channel environment. The simulation result of the artificial potential field method and autonomous path planning method are shown in Fig. 9 a and b respectively. The experiment result of double U environment is shown on top figure, and the experiment result of narrow channel environment is shown on bottom figure. The simulation results showed that the robot is inclined to be trapped in a point by using an artificial potential field method in U obstacles environment, while it will get rid of the local minimum point difficultly and finally will be trapped in the point by using autonomous learning strategy. Otherwise, the robot appears fluctuating by using artificial potential field method in narrow channel environment, while it can learn the optimal short path with few swings by using autonomous learning strategy.
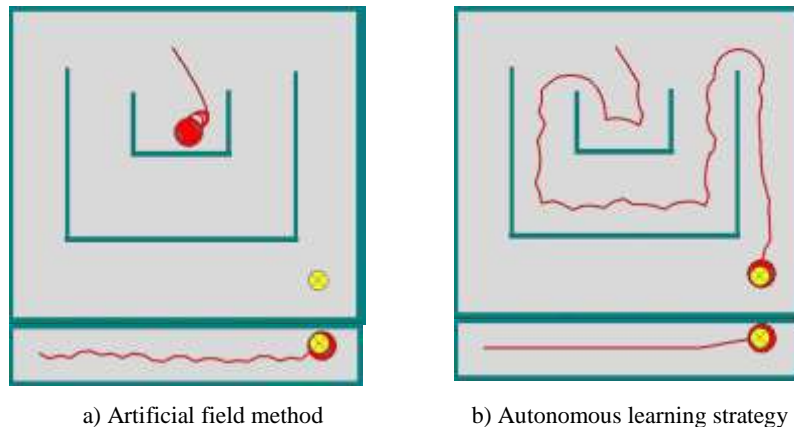

a) Artificial field method   b) Autonomous learning strategy

Fig. 9. Comparison experiments of overcoming the local minimum and oscillation

5. Conclusion

In the paper, autonomous learning strategy which simulates operant conditioning mechanism, is put forward to realize the robot autonomous path planning. The path planning simulation of unknown environment is executed in Mobotsim environment and the simulation results prove that:

1) Robot can successfully evade obstacles and arrive at goal point with optimal path based on bionic autonomous learning strategy.

2) The designed autonomous learning strategy effectively solves the local minima problem in special environment and also left and right oscillation problem in narrow environment, so it is better than artificial field method.

References

1. C a i, Z.-X., H.-G. H e, C. H o n g. Mobile Robot Navigation Control Theory and Method in Unknown Environment. Science Press, 2009.
2. Ko r e n, Y., J. B o r e n s t e i n. Potential Field Methods and Their Inherent Limitations for Mobile Robot Navigation. – In: Proc. of 1991 IEEE International Conference on Robotics and Automation (Cat. No 91CH2969-4), 9-11 April 1991, Los Alamitos, CA, USA, IEEE Comput. Soc. Press, 1991, pp. 1398-1404.
3. Q i a o, J. F., R. Y. F a n, H. G. H a n, X. G. R u a n. Research and Realization of Dynamic Neural Network Navigation Algorithm for Mobile Robot. – Control Theory & Applications, Vol. **27**, January 2010, No 1, pp. 111-115.
4. C h a o m i n, L., S. X. Y a n g. A Bioinspired Neural Network for Real-Time Concurrent Map Building and Complete Coverage Robot Navigation in Unknown Environments. – IEEE Transactions on Neural Networks, Vol. **19**, 2008, No 7, pp. 1279-1298.
5. C a i h o n g, L., Z. J i n g y u a n, L. Y i b i n. Application of Artificial Neural Network Based on Q-Learning for Mobile Robot Path Planning. – In: Proc. of 2006 IEEE International Conference on Information Acquisition, 20-23 August 2006, Piscataway, NJ, USA, IEEE, 2006, pp. 978-982.
6. Q i a o, J., R. F a n, H. H a n et al. Q-Learning Based on Dynamical Structure Neural Network for Robot Navigation in Unknown Environment. – In: Proc. of 6th International Symposium on Neural Networks, ISNN 2009, 26-29 May 2009, Wuhan, China, Springer Verlag, 2009, pp. 188-196.
7. D a l a l a h, D., M. T. H a y a j n e h. Dynamic Neural Model for Mobile Robot Navigation. – Neural Network World, Vol. **17**, 2007, No 1, pp. 65-80.
8. W a n g, X.-S., Y. G a o, Y.-H. C e n g. Knowledge Guided Genetical Gorithmfor Path Planning of Robot. – Control and Decision Making, Vol. **24**, 2009, No 7, pp. 1043-1049.
9. T z a f e s t a s, C. S., S. G. T z a f e s t a s. Recent Algorithms for Fuzzy and Neurofuzzy Path Planning and Navigation of Autonomous Mobile Robots. – Systems Science, Vol. **25**, 1999, No 2, pp. 25-39.
10. C h e n, W.-D., B.-X. L i, Q.-G. Z h u. Application of Fuzzy Control in Mobile Robot Path Planning. – Computer Engineering and Applications, Vol. **45**, 2009, No 31, pp. 221-223.
11. Y u h u, C., W. X u e s o n g, L. R u h a i. Fuzzy Control System for Path Following of Mobile Robots. – ICIC Express Letters, Vol. **3**, 2009, No 3A, pp. 403-408.

12. Z a v l a n g a s, P. G., S. G. T z a f e s t a s. Motion Control for Mobile Robot Obstacle Avoidance and Navigation: A Fuzzy Logic-Based Approach. – Systems Analysis Modelling Simulation, Vol. **43**, 2003, No 12, pp. 1625-1637.
13. W a n g, X.-S., Y.-H. C e n g. Machine Learning Theory, Method and Application. Science Press, 2009.
14. X u, X i. Reforment Learning and Approximate Dynamic Programming. Science Press, 2010.
15. Y e, C., N. H. C. Y u n g, D. W a n g. A Fuzzy Controller with Supervised Learning Assisted Reinforcement Learning Algorithm for Obstacle Avoidance. – Systems, Man, and Cybernetics, Part B: Cybernetics. – IEEE Transactions on, Vol. **33**, 2003, No 1, pp. 17-27.
16. B o u b e r t a k h, H., M. T a d j i n e, P. Y. G l o r e n n e c. A New Mobile Robot Navigation Method Using Fuzzy Logic and a Modified Q-Learning Algorithm. – Journal of Intelligent and Fuzzy Systems, Vol. **21**, 2010, No 1, pp. 113-119.
17. C h a n g, C., P. G a u d i a n o. Application of Biological Learning Theories to Mobile Robot Avoidance and Approach Behaviors. – Journal of Complex Systems, Vol. **1**, 1998, pp. 79-114.