



Path Planning Based on Q-learning and Three-segment Method for Aircraft Fuel Tank Inspection Robot

Niu Guochen^a, Xu Kailu^b

^aRobotics Institute, Beihang University, Beijing 100191, China

^bRobotics Institute, Civil Aviation University of China, Tianjin 300300, China

Abstract. In order to realize the path planning of continuum robot for inspecting defects in the aircraft fuel tank compartment, an approach based on Q-learning and Three-segment Method was proposed, and the posture of the robot meeting the inherent and spatial structure constraint requirements was planned. Firstly, the simulation model of the aircraft fuel tank was established. Moreover, the workspace was rasterized to decrease the computing complexity. Secondly, the Q-learning algorithm was applied and the path from the initial point to the target was generated. In terms of target guided angle and three-segment method, the joint variables corresponding to each transition point on the path could be obtained. Finally, the robot reached the target by progressively updating the joint variables. Simulation experiments were implemented, and the results verified the effectiveness and feasibility of the algorithm.

1. Introduction

In civil aviation aircraft maintenance, the fuel tank check is usually performed manually to find out the leak or corrosion location, and the crew have to enter the tank to check it. It is a cumbersome and time-consuming work, while the environment with heavy concentration oil and gas is harmful to human health. Hence, an automatic device is needed to help the crew to do the work. Comparing with the traditional discrete robot, the continuum robot with redundant degrees of freedom can reach places usually unachievable for rigid link robots. So our project group designs an Aircraft Fuel Tank Inspection Robot (AFTIR) with continuous structure to assist the crew to inspect the fuel tank [1].

To perform the space motion of a continuum robot, it is necessary to plan the path of the robot first. Several researchers have proposed methods about continuum robot path planning. The traditional methods for manipulator path planning include rapid expansion of random tree [2-3], polynomial interpolation [4], artificial potential field method [5], probability roadmap method [6], machine learning [7-8], etc. In reference [9], the authors evaluated four trajectory generation strategies in terms of the resulting Cartesian paths and spatial extent of the course of motion. A points-based safe path planning (PoPP) algorithm for continuum robots was proposed in reference [10]. The authors exploited the constant curvature-bending property of continuum robots in their path planning process. The algorithm is computationally efficient and provides a good tradeoff between accuracy and efficiency. This algorithm also provides information regarding path

2010 *Mathematics Subject Classification.* 68T05.

Keywords. Three-segment Method; Q-learning, continuum robot; Path planning, Rasterizing.

Received: 23 October 2017; Accepted: 30 January 2018

Communicated by Hari M. Srivastava

Email addresses: niu.guochen@139.com (Niu Guochen), 799673781@qq.com (Xu Kailu)

volume and flexibility in movement. However, this algorithm only considers the planar path planning, it is difficult to apply to the 3D space. In reference [11], the authors proposed a learning from demonstration method into the motion planning of ionic polymer-metal composite IPMC based flexible and soft robotic manipulators for their tasks such as crossing a hole. The demonstration paths were encoded by Gaussian mixture model, and recommended paths were generated by Gaussian mixture regression. In reference [12], the authors calculated the joint variables based on the combination of newton method and intestinal mathematical model to solve the problem of intestinal endoscope robot path planning. But it is difficult to extend to multi-joint control. An obstacle avoidance algorithm for OctArm [13] robots was presented in reference [14], and this algorithm was used to build and solve the kinematic model of the plane bending characteristic. Meanwhile the real-time adaptive motion planner was designed based on the plane curve parameters. But its kinematic model is not suitable for other types of robots. In reference [15], the authors presented a pose estimation and obstacle avoidance approach for tend-driven multi-segments continuum manipulators moving in dynamic environment. The strategy will guide the manipulators tip to desired target and keep the manipulator body avoid obstacles. A search method based on region clipping deduced from continuity analysis of the robot end position was presented in reference [16]. Minimum Distance Summation (MDS) was adopted to select the relatively optimal path as the evaluation criterion. This method effectively reduces the computational complexity of path searching. In reference [17-18], a 3D path tracking approach based on end point approximation and a closed-loop fuzzy controller based on attitude feedback is proposed to track the planned path and reach the target for after during the maintenance process of aircraft fuel tank. In reference [19], an algorithm based on Target Guided Angle (TGA) was presented to realize path planning of AFTIR. Strategies of dimension reduction and regional division were proposed to decrease the high time complexity of blind search. Then the search process was optimized and evaluation function was proposed to evaluate results of search algorithm, but the algorithm is only suitable for single compartment environment of the aircraft fuel tank and the planned path is in a plane. In fact, spatial path are necessary for the robot to inspect the aircraft fuel tank.

The operation space of the continuum robot is three-dimensional space, which is restricted by its own high redundancy structure and joint coupling. Its path planning and collision avoidance problems are more complicated than mobile robots [20]. Due to the difficulty of solving inverse kinematics, simply using the Cartesian path planning method will make the workspace traversal and the multipoint collision avoidance of the robot become complicated. On the basis of the research on the algorithm based on TGA, an approach based on Q-learning and the TSM is proposed in this paper to get the path of multi-compartment environment. The path of the robot end point is calculated through Q-learning and the joint segment variables are calculated according to the TGA and TSM. The path satisfied with the external environment constraint, and robot structure constraint can be obtained.

2. The raster of workspace

The basic idea of the approach mentioned in this paper is as follows. Firstly, the workspace is rasterized to decrease computing time complexity. The Q-learning algorithm is applied to obtain the path between the initial point and the target. The path consists of the initial raster, intermediate raster named after transition states and terminal raster. Secondly, the joint variables corresponding to each transition state will be obtained through TGA and TSM. Thirdly, the robot can reach the terminal raster by progressively updating the joint variables. Finally, the joint variables are fine-tuned when the robot reaches the terminal raster. The control flow schematic diagram of AFTIR is shown in Figure.1. As shown in Figure 2, the structure of AFTIR mainly consists of three parts including the detector, the snake arm and the linear module. The detector perceps the environment using the endoscope. The snake arm enters the interior of the fuel tank for inspection, and the rigid linear module provides telescopic movement of snake arm. The snake arm consists of several continuum sections. Each section consists of some support disks, driving cables and a flexible backbone, and it can bend and rotate through changing lengths of three driving cables which are uniformly distributed around backbone at intervals of 120° . Figure.3-(a) shows the internal environment of the aircraft fuel tank in which some oil pipes and stringers are distributed. In order to compute the path conveniently, the system frame S is established with its origin located at the center of the access panel. The

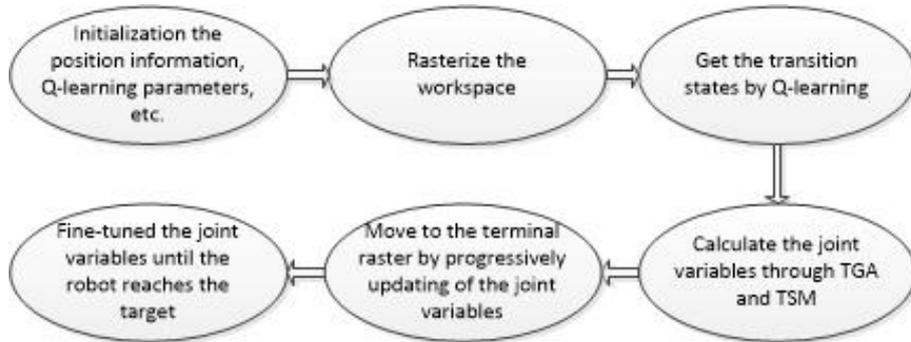


Figure 1: Control flow schematic diagram of AFTIR

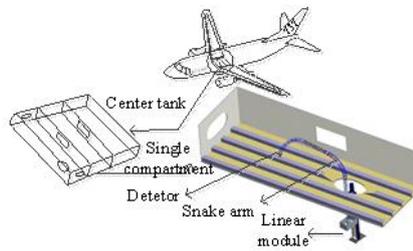
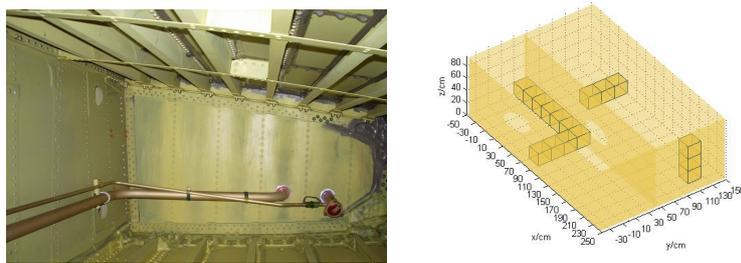


Figure 2: Schematic diagram of AFTIR

Z_S axis is perpendicular to the access panel, the X_S axis is coincident with the minor axis of the access panel. The workspace is divided into $N_x * N_y * N_z$ equal square raster. Each raster corresponds to a serial number $(a, b, c), a \in \{1, 2, \dots, N_x\}, b \in \{1, 2, \dots, N_y\}, c \in \{1, 2, \dots, N_z\}$.



(a) Aircraft fuel tank real environment (b) Simulation mode of aircraft fuel tank

Figure 3: The model of aircraft fuel tank

The obstructed space is defined as the raster passing through the tank baffle, the tank crust and other obstacles, and its complementary set is the free space of the robot. As shown in Figure.3-(b), the shadow is obstructed space. Then, the path from the initial point to the target will pass through a series of raster which do not belong to obstructed space. The path is represented as a finite number of center points of the raster. That is $Path = \{S, P_1, P_2, \dots, P_i, \dots, T\}$. Where, $i = 1, 2, \dots$.

3. Path planning based on Q-learning

The Q-learning algorithm is a kind of reinforcement learning algorithm based on a time difference strategy. Reinforcement learning holds the promise of enabling autonomous robots to learn large repertoires of behavioral skills with minimal human intervention[21]. It refers to an expected utility function after a certain action in a given state. The principle of the reinforcement learning is based on learning by trial and error. The algorithm does not require accurate modeling of the system, nor does it require a large amount of training data [22].

On each step of interaction, the controller receives an input that provides indications of the current state. Then, the controller chooses an action a which will change the state of the system, and it receives a reward according to the new state. The task of the controller is to find a policy, mapping actions to states, that maximizes the long-run sum of rewards. In this paper, the task of the robot for Q-learning is to find an optimal sequence of actions $\pi^* (s_i = a)$, that is, the cumulative return value obtained from the initial state s_i to the target is the largest. In this article, Optimization means the least number of actions. The basic idea of Q-learning is to evaluate the utility function Q and the optimal strategy on the basis of the return sequence. Knowledge can be acquired from lots of tries. Once a reward is obtained, an agent can update its previous knowledge (Q). The training is shown as formula (1).

$$\begin{cases} Q(s, a) = r(s, a) + \gamma \cdot \max_{a'} Q(\delta(s, a), a') \\ \pi^*(s) = \arg \max_a Q(s, a) \\ \hat{Q}(s, a) \leftarrow r(s, a) + \gamma \cdot \max_{a'} \hat{Q}(s', a') \end{cases} \quad (1)$$

Where, s is a state. a is an action. γ is a discount factor. (s, a) and a' are the next state and the next action. $\pi^*(s)$ is the optimal action choice for the current state. $r(s, a)$ is the immediate reward value after performing action a in state s . As the simplicity of the state space in this paper, its initial formula is (2). T is the target space. O is the obstacle space. F is the freedom space. According to the recursive definition of the Q function, the basic rules of iterative approximation to the Q algorithm are obtained.

$$r(s, a) = \begin{cases} 100, & s \in T \\ -100, & s \in O \\ 0, & s \in F \end{cases} \quad (2)$$

In order to ensure the balance between the exploration of the unknown path and the experience which has been obtained, the ϵ -greed method is adopted. The probability of each state to be explored is ϵ , and the probability of being developed is $1-\epsilon$. Once a certain action is chosen, the action scores will be updated. With the continuous deepening of study, ϵ will become more and more small. The mode of learning turns from full exploration to intensive study.

After the workspace is rasterized, the raster at which the end of the continuum robot located is defined as the state in the Q table, and the nearest six movements to the raster around are defined as six actions of the Q table. Then this system is modeled as a definite Markov process. The return function is bounded. And the action selection mechanism allows each state action to be infinitely accessed. Thus, the optimal sequence of actions from the initial point to the target can be obtained after lots of training. As shown in Figure 4. The pentagram represents the target. The diamond represents the initial point. The circles represent the transition point on the path.

4. Posture generation based on three-segment method

In order to reach every transition point, the posture generation method of the robot is required. To obtain the configuration parameters of the robot moving in a single compartment, a posture generation algorithm based on TGA was presented in reference [19]. The high redundancy characteristic of continuum structure leads to the high time complexity of path searching in 3D space. Dimension reduction strategy was adopted, the 3D path planning is converted to the 2D path planning through introducing the reference

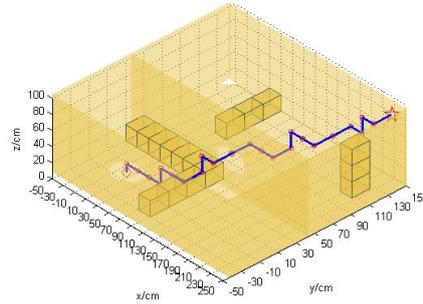


Figure 4: Transition points obtained by Q-learning

plane. The search for the rotation angle is eliminated. The search range of the joint bending angle is greatly reduced due to the concept of guided angle. In this way, the search time complexity is greatly reduced. The concept of reference straight line is also proposed. The distance between the searched path and the reference line is taken as the evaluation criterion of the planning posture. The posture generated in this way has the following characteristics: 1. The planning posture is in a reference plane. 2. If the two positions of the target are close in distance, the two postures of the robot are also similar. That is, if the two points of space attitude transformation are very close, the amount of joint variable changes will be small when the posture is updated.

There are multiple compartments in the aircraft fuel tank which are connected by internal hank hatchways. As shown in Figure 5, two compartments of the center fuel tank are represented by 1# and 2# respectively. A path from the inlet of the fuel tank to the point G is planned. The planar posture in single

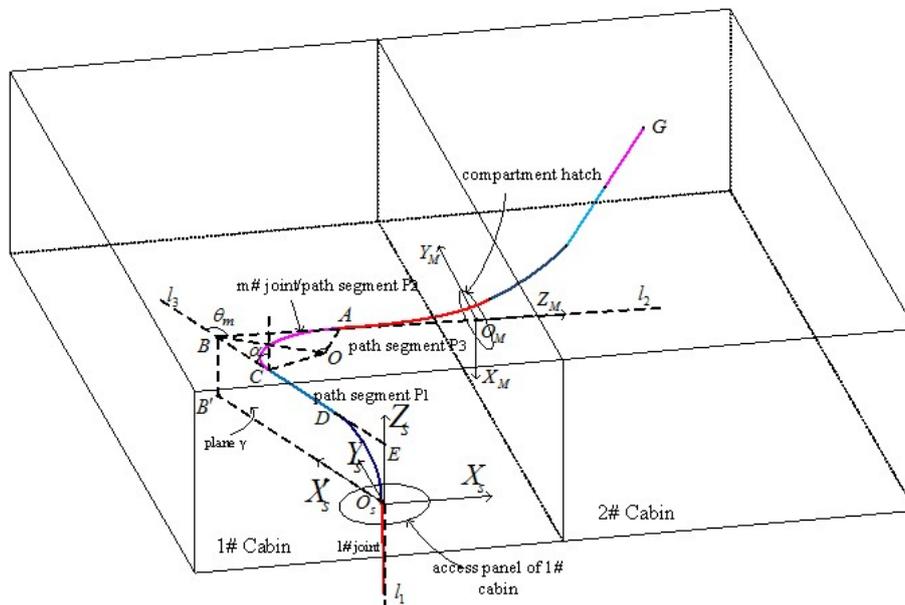


Figure 5: Schematic diagram of TSM

compartment can be obtained by TGA, but the posture through multiple compartments can not be obtained, because it is inevitable to plan the spatial posture. On the basis of the research on TGA, TSM is designed. The posture is decomposed into three parts, two of which (AG, O_xC) are obtained by TGA and another section (AC) is obtained by geometric analysis. The three parts of the posture are generated respectively.

Table 1: Nomenclature used in this paper

Symbol	Description
{S}	The system frame {S} is attached to the aircraft fuel tank access panel with its origin located at the center of the access panel. The Z_S axis is perpendicular to the access panel, X_S axis is coincident with the minor axis of the access panel
{M}	Frame {M} is attached to the compartment hatch between the compartment 1# and 2#. Its origin O_M is located at the center of the 2# compartment hatch. The X_M axis points to the compartment bottom. The Z_M axis points to 2# compartment along the mid-perpendicular of the compartment hatch.
l_1	The mid-perpendicular of the tank inlet.
l_2	The mid-perpendicular of the compartment hatch, the tangent of the head end of the P3 part.
l_3	The tangent of the terminal end (DC) of the P1 part.
γ	The plane determined by l_1 and l_3 .
A,G	Two ends of P3.
A,C	Two ends of P2.
O_S, C	Two ends of P1.
P1	The first part of the posture.
P2	The second part of the posture.
P3	The third part of the posture.
α	The angle of l_3 and Z_S axis.
δ	The angle of γ and $X_S O_S Z_S$.
d	$ BC = BA = d$
l	Length of the single segment.
θ_m	Bending angle of P2 part.

Then they are combined together to get the whole compartment posture.

The detailed process of the algorithm is as follows:

Step 1. Frame {M} relative to {S} can be described as formula (3).

$${}^S_M T = \begin{bmatrix} {}^S_M R & {}^S P_{O_M} \\ 0 & 1 \end{bmatrix} \tag{3}$$

Then,

$${}^M P_G = {}^S_M T^{-1} \cdot {}^S P_G \tag{4}$$

Where, ${}^S P_{O_M}$ is the position vector of O_M relative to {S}, ${}^S P_G$ and ${}^M P_G$ are position vectors of target point G relative to {S} and {M} respectively. The posture in 2# compartment P3 is planned in frame {M} according to TGA. The posture parameters include the joint number N_{p3} , the ascent distance dis_{p3} , the joint bending angles and the joint rotation angles. The coordinates of the point A is shown in (5).

$${}^S P_A = ({}^S x_A, {}^S y_A, {}^S z_A, 1)^T = ({}^S x_{O_M} - h, {}^S y_{O_M}, {}^S z_{O_M}, 1)^T \tag{5}$$

Where, $h = N_{p3} \cdot l - dis_{p3}$.

Step 2. Derivation of P_1 .

To compute parameters of P_1 in 1# compartment, coordinates of the space points B, E and C should be determined successively through geometrical analysis method. Make $|BC| = |BA| = d$, in terms of the geometrical relationship, the formula (6) can be obtained.

$$\tan(\theta_m/2) = d/(l/\theta_m) \tag{6}$$

The coordinates of B is shown in (7).

$${}^S P_R = ({}^S x_R, {}^S y_R, {}^S z_R, 1)^T = ({}^S x_A - h, {}^S y_A, {}^S z_A, 1)^T \tag{7}$$

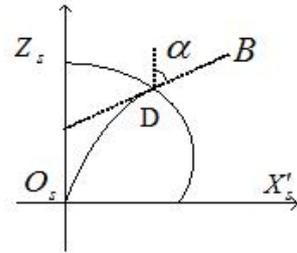


Figure 6: Robot posture with the target point of B

In frame $\{S\}$, the target orientation is α when we set B as the target according TGA, as shown in Figure.6, here $|DB| > d$.

In plane γ , supposing the equation of l_3 which goes through B and has an included angle α with Z_s axis. Supposing $z = kx + b$, where $k = \cot(\alpha)$, and $b = {}^S z_B - \cot(\alpha) \cdot \sqrt{({}^S x_B)^2 + ({}^S y_B)^2}$, Then the coordinates of E can be given by (8).

$${}^S P_E = (0, 0, {}^S z_E, 1)^T = (0, 0, {}^S z_B - \cot(\alpha) \cdot \sqrt{({}^S x_B)^2 + ({}^S y_B)^2}, 1)^T \tag{8}$$

Obviously, we have (9) and (10).

$$\theta_m = \arccos \frac{{}^S x_B}{\sqrt{({}^S x_B)^2 + ({}^S y_B)^2 + ({}^S z_B - {}^S z_E)^2}} \tag{9}$$

$$\tan(\delta) = {}^S y_B / {}^S x_B \tag{10}$$

d, θ_m, δ and ${}^S P_B$ can be obtained by the above formula. Finally, coordinate of C is described as (11). Using TGA, the parameters of P_1 will be available by choosing C as the target point.

$${}^S P_C = ({}^S x_C, {}^S y_C, {}^S z_C, 1)^T = \begin{cases} (\sqrt{({}^S x_B)^2 + ({}^S y_B)^2} - d \cdot \sin \alpha) \cdot \cos \delta \\ (\sqrt{({}^S x_B)^2 + ({}^S y_B)^2} - d \cdot \sin \alpha) \cdot \sin \delta \\ {}^S z_B - d \cdot \cos \alpha \\ 1 \end{cases} \tag{11}$$

Step 3. Derivation of P_2

From analysis above, we can get that C is the origin of frame $\{m - 1\}$. According to the joint variables of the first $m-1$ joint segments and the multi joint kinematics model, the homogeneous transformation matrix of frame $\{m - 1\}$ to $\{S\}$ is ${}^S_{m-1} T$. So the coordinates of A in frame $\{m - 1\}$ can be described as (12).

$${}^{m-1} P_A = {}^S_{m-1} T^{-1} \cdot {}^S P_A \tag{12}$$

P_2 is the curvature consisting of $m\#$ joint segment. The bending angle θ_m has been obtained in step 2. The rotation angle φ_m is calculated according to the kinematic model of single segment in frame $\{m - 1\}$, and the formula (13) can be obtained.

$$\varphi_m = \arctan({}^{m-1} y_A / {}^{m-1} x_A) \tag{13}$$

Step 4. Summarize parameters of the planning path.

After coordinate transformation and posture combination, the parameters of the whole posture can be obtained. As in formula (14).

$$\begin{cases} N = N_{p1} + 1 + N_{p3} \\ dis = (N_{p3} + 1) \cdot l + dis_{p1} \\ \theta = \{\theta_1, \theta_2 \cdots, \theta_{m-1}, \theta_m, \theta_{m+1}, \cdots, \theta_{m+N_{p3}}\} \\ \varphi = \{\varphi_1, \varphi_2 \cdots, \varphi_{m-1}, \varphi_m, \varphi_{m+1}, \cdots, \varphi_{m+N_{p3}}\} \end{cases} \tag{14}$$

Where, $N_{p1} = m - 1$ is the joint number of P_1 . N_{p3} is the joint number of P_3 . N is the total number of joints. dis and dis_{P1} represent the ascent distance of the whole path and the P_1 respectively. θ and φ represent the bending and the rotation angle set of the whole posture respectively.

5. Simulation results and analysis

The iteration times using blind search strategy is k in formula (15). Where, N is the joint number of the posture. $\Delta\theta$ and $\Delta\varphi$ represent the step size in search of the bending angle and the rotation angle respectively. The range of rotation angle and bending angle are $[0, 2\pi]$ and $[0, \pi]$. The iteration times using the method proposed in this paper is given by formula (16). Obviously, the TSM has great advantages in computing time complexity.

$$k = \left(\frac{\pi}{\Delta\theta}\right)^N \cdot \left(\frac{2\pi}{\Delta\varphi}\right)^N \tag{15}$$

$$k = 3 \times \frac{\pi}{\Delta\theta} \tag{16}$$

Considering the mechanical structure and specific dimension of the fuel tank of Boeing 737-200, the center fuel tank model consisting of two compartments is established. Four different target points are chosen to verify the TSM, and the result are shown in Figure 7 (a), (b), (c) and (d) respectively. The simulation space was divided into $15 \times 10 \times 5$ raster. Simulation experiments are performed on a 3.20 GHz Intel(R)Core(TM) PC with 8 GB RAM. The length of single joint is 50 cm.

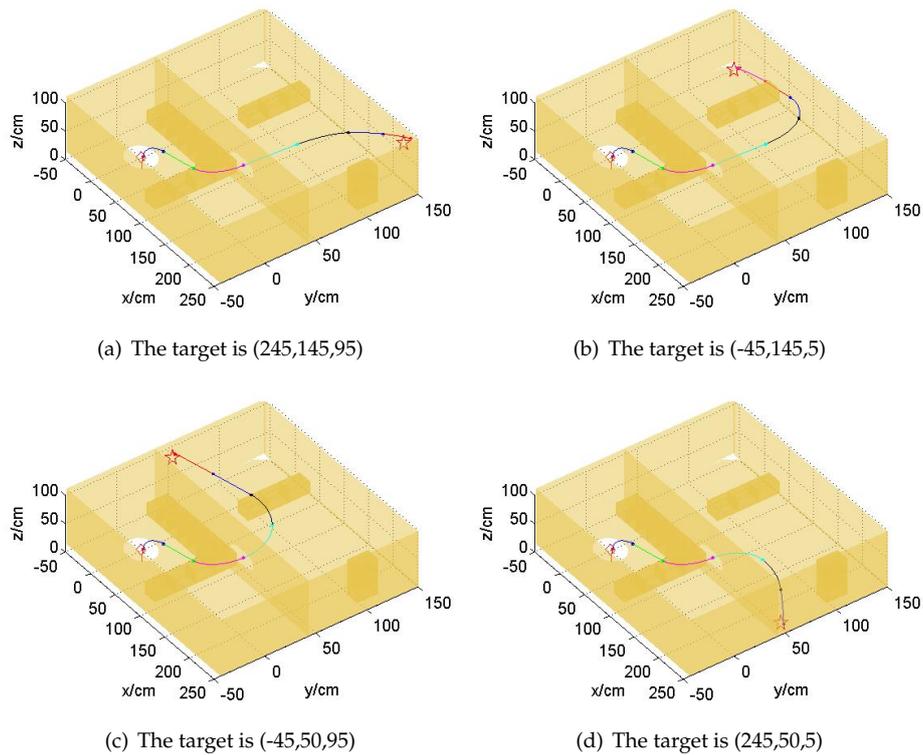


Figure 7: Posture generation of four targets

Experimental results corresponding to four targets are shown in Table 2. Where the desired position represents the given coordinates of the target. The actual position is the coordinates which the robot end reaches, and it is calculated through the forward kinematics of the continuum robot. Dis represents the

Table 2: Experiment data

	a	b	c	d
Desired position(cm)	(245, 145, 95)	(-45, 145, 5)	(-45, 50, 95)	(245, 50, 5)
Actual position(cm)	(243.8,146.5, 93.2)	(-43.2,144.9,5.8)	(-42.3, 52, 94.2)	(246.1, 49.5, 4.8)
operation time t(s)	2.441	3.452	2.427	1.853
(cm)	1.33	2.45	1.11	0.82
N	8	9	8	7
Dis(cm)	352.58	402.58	351.51	302.59
$\varphi_1, \varphi_2, \dots, \varphi_N$ (deg)	(6.16,6.16,6.16,95.29, -106.70,-106.70, -106.70,-106.70)	(6.16,6.16,6.16,95.29, 87.90,87.90,87.90, 87.90,87.90)	(6.16,6.16,6.16,95.29, 117.85,117.85,117.85, 117.85)	(6.16,6.16,6.16,95.29, -66.25,-66.25,-66.25)
$\theta_1, \theta_2, \dots, \theta_N$ (deg)	(7.26,79.26,0.06,79.69, ?5.40,61.20,10.80,5.40)	(7.26,79.26,0.06,79.69, 0,61.20,32.40,0,10.80)	(7.26,79.26,0.06,79.69, ?72.00,39.60,0,1.80)	(7.26,79.26,0.06,79.6 9,77.40,28.80,3.60)

ascent distance of the base. t represents the operation time. N represents the joint number that the current posture needs. $\varphi_1, \varphi_2, \dots, \varphi_N$ and $\theta_1, \theta_2, \dots, \theta_N$ denote the rotation angle set and bending angle set of the robot respectively. σ represents the Euclidean distance between the desired position and the actual position. The maximum deviation is about 2.45 cm, which meets the needs of the project. The main reason for the error is the selection of the step size in search. The longest operation time of the algorithm is 3.45s as the robot using nine joint segments. The program runs for less than two seconds in most cases. It is obvious that the TSM greatly reduces the computation complexity. The maximum joint number is up to nine to reach the vertices of the fuel tank. That is, in order to achieve the traversal check of the two compartments, the continuum robot requires at least nine joints.

Figure 8 illustrates the posture transformation process of the robot. During the transformation process, the terminal raster was determined according to the input target coordinates (the pentagram), and the transition points (the red point on the polyline) of the robot were obtained by Q-learning. Each transition posture was obtained by TGA and TSM, as shown in Figure 7 using the dashed line (here only six transition postures are shown, which is used to describe the movement of the robot to the target).The robot approach the terminal raster by progressively updating of the posture. When reaching the terminal raster, the TSM was used again to obtain the posture to the target, as shown in figure 7 using the solid line.

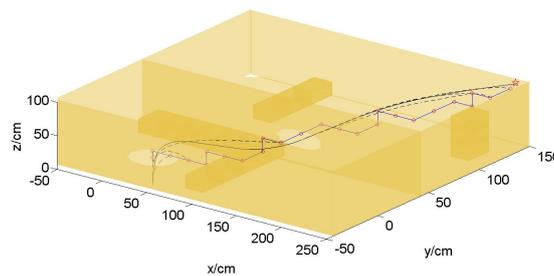


Figure 8: Posture transformation process of AFTIP

6. Conclusion

In this paper, an approach about path planning and posture generating based on Q-learning and three-segment method was proposed to reach the target for AFTIR during the maintenance process of aircraft fuel

tank. On the basis of spatial discretization, the Q-learning algorithm was used to get the path of the robot end point. Based on TGA, which was previously proposed by our studying team, the TSM was designed to generate the spatial posture of the continuum robot to reach the target. Then, along the transition point on the path, the robot reaches each raster by updating the joint variables in turn. The transformation of the spatial posture was realized. Simulations of the algorithm were carried out. The results showed the feasibility and effectiveness of the algorithm. The robot can reach every point in the searching space using the algorithm.

7. Acknowledgements

This work was funded by the Tianjin Research Program of Application Foundation and Advanced Technology #14JCQNJC04400. The authors would like to thank their colleagues on the robotics institute of Beihang University and Civil Aviation University of China for their significant contributions.

References

- [1] Gao Qingji, Wang Weijuan, Niu Guochen, Wang Lei, Zheng Zunchao (2013). "Study of bionic structure and kinematics of robot for aircraft fuel tank inspection." *Acta Aeronautica*. vol.34, No.7, pp.1748-1756.
- [2] Wu Keyu, Wu Liaa, Ren Hongliang (2015). "Motion planning of continuum tubular robots based on centerlines extracted from statistical atlas." *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*. pp. 5512-5517.
- [3] Ding Runwei, Ma Jiaxiao, Kang Rensheng, Zhong Chengcheng (2015). "Region classification based robot security path planning in dynamic environment." *Journal of Huazhong University of Science and Technology*. vol.43 pp.312-314.
- [4] Jiang Wei, Wu Gongping, Wang Wei, Zhang Jie (2016). Manipulator dynamic modeling and motion planning for working robot. *Chinese Journal of Engineering*. vol.38, No.6, pp. 867-875.
- [5] Wei, Zhixuan, Chen Weidong, Wang Hesheng (2017). "Shared Control with Flexible Obstacle Avoidance for Manipulator." *Advances in Intelligent Systems and Computing*. vol.531, pp.229-241.
- [6] Yu Xiaowen, Zhao Yu, Wang Cong (2017). "Trajectory planning for robot manipulators considering kinematic constraints using probabilistic roadmap approach." *Journal of Dynamic Systems Measurement and Control-Transactions of the Asme*. vol.139, No.2.
- [7] Thuruthel, Thomas George, et al (2016). "Learning Global Inverse Statics Solution for a Redundant Soft Robot." *Int. Conf. on Informatics in Control, Automation and Robotics*. pp.303-310.
- [8] Reiter, Austin, et al (2011). "A learning algorithm for visual pose estimation of continuum robots." *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*. pp.2390-2396.
- [9] Fellmann Carolin, Burgner-Kahrs Jessica (2015). "Implications of trajectory generation strategies for tubular continuum robots." *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*. pp.202-208.
- [10] Shahzad Khuram, Iqbal Sohail, Bloodsworth Peter (2015). "Points-based Safe Path Planning of Continuum Robots." *International Journal of Advanced Robotic Systems*. vol.12.
- [11] Wang, Hongqiang, Chen Jie, Lau Henry, Ren Hongliang (2016). "Motion Planning Based on Learning From Demonstration for Multiple-Segment Flexible Soft Robots Actuated by Electroactive Polymers." *IEEE Robotics and Automation Letters*. vol.1, No.1, pp.391-398.
- [12] Hu Haiyan, Wang Pengfei, Sun Lining, Zhao Bo, Li Mantian (2010). "Kinematic Analysis and Simulation for Cable-driven Continuum Robot." *Journal of Mechanical Engineering*. vol.16, No.19, pp.1-8.
- [13] Pritts Michael B. Rahn Christopher D.(2004). "Design of an artificial muscle continuum robot." *IEEE Int.Conf.on Robotics and Automation*. pp. 4742-4746.
- [14] Xiao, Jing, Vatcha Rayomand (2010). "Real-time adaptive motion planning for a continuum manipulator." *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*. pp.5919-5926.
- [15] Ataka, Ahmad, et al (2016). "Real-Time Pose Estimation and Obstacle Avoidance for Multi-segment Continuum Manipulator in Dynamic Environments." *IEEE Int. Conf. on Intelligent Robots and System*. pp.2827-2832.
- [16] Niu, Guochen, Zheng Zunchao, Gao Qingji (2014). "Collision free path planning based on region clipping for aircraft fuel tank inspection robot." *IEEE Int. Conf. on Robotics and Automation*. pp.3227-3233.
- [17] Niu Guochen, Wang Li, Gao Qingji, Hu Dandan (2014). "Path-tracking Algorithm for Aircraft Fuel Tank Inspection Robots." *International Journal of Advanced Robotic Systems*. vol.11, No.1.
- [18] Niu Guochen, Wang Li, Zong Guanghua (2015). "Attitude control based on fuzzy logic for continuum aircraft fuel tank inspection robot." *Journal of Intelligent and Fuzzy Systems*. vol.29, No.6, pp.2495-2503.
- [19] Gao, Qingji, Wang Lei, Niu Guochen, Wang Weijuan (2013). "Path planning for continuum robot based on target guided angle." *Journal of Beijing University of Aeronautics and Astronautics*. vol.39, No.11, pp.1486-1490.
- [20] Qian Kun, Jia Kai, Song Xiao (2015). "Robot manipulator avoidance planning based on low-dimensional mapping and Q-learning." *Journal of Huazhong University of Science and Technology*. vol.43, pp.468-472.
- [21] Gu, Shixiang, Holly, Ethan, Lillicrap, et al(2016). "Deep Reinforcement Learning for Robotic Manipulation with Asynchronous Off-Policy Updates." *IEEE Int. Conf. on Robotics and Automation*. pp. 3389-3396.
- [22] Zhao hui, Liu Yazhe (2016). "Application of improved Q learning algorithm in trajectory planning." *Journal of Jilin University (Information Science Edition)*. vol.34, pp.468-472.