Transactions on Computational and Scientific Methods | Vo. 4, No. 2, 2024
ISSN: 2998-8780
https://pspress.org/index.php/tcsm
Pinnacle Science Press

Enhancing Path Planning Efficiency in Virtual Assembly Environments Using Fuzzy Bayesian Deep Q-Networks

Rachel Young

JohnsHopkinsUniversity, Baltimore, USA

Ryry909@jhu.edu

Abstract: The utilization of computer virtual reality technology in virtual assembly can facilitate the design and planning of the assembly process, enhance production efficiency, and decrease economic costs. Path planning represents a critical area of development within virtual assembly, making the investigation of path planning technology essential for designing assembly paths in intricate environments. This paper introduces a deep Q-network algorithm grounded in Fuzzy Bayes, which has been evaluated within a virtual assembly context. The findings indicate that the Fuzzy Bayes-based deep Q-network algorithm demonstrates superior navigability and planning efficiency in complex, confined environments.

Keywords: Path planning; deep reinforcement learning; fuzzy Bayesian decision algorithm; fuzzyBayesian deep Q network.

1. Preface

This year, With the rapid development of virtual reality technology, it provides a new solution for industrial simulation assembly technology.Virtual reality technology is applied to the assembly design of mechanical and electrical products, and the virtual assembly technology is born. Virtual assembly is an important part of virtual manufacturing, By using computer to simulate the real assembly operation and demonstrate the assembly process with visual tools, problems can be found as soon as possible and a more reasonable assembly scheme can be formulated. This method saves time and effort, improves performance, easily finds problems, and can be modified in time. It greatly shortens the R & D and manufacturing cycle of products, reduces the cost of products and improves the competitiveness of products.In virtual assembly, assembly path planning technology can realize the automatic generation of path, which can greatly reduce the economic cost. Therefore, the research of automatic path planning technology has practical significance for virtual assembly.In this paper, we use reference [1] to find the shortest path of the document; In reference [2], an improved A* algorithm is proposedto solve the robot path planning problem, and the experimental results show that the success rate of the improved algorithm is higher than that of the original A* algorithm.

Reference [3] searched randomly assigned nodes in the space, and calculated and selected the shortest route.In reference [4], a new particle swarm optimization algorithm based on non- uniform Markov chain is constructed, which provides a new method for intelligent robot path planning.In reference [5], the path planning ability of robot is trained by using full convolution neural network and fast extended random tree algorithm.Reference [6] combines human- computer interaction guidance and path solving algorithm to guide the generation of assembly path.In reference [7], the artificial potential field local optimization algorithm is introduced into ant colony algorithm to improve the convergence speed of global path planning.In reference [8], an improved particle swarm optimization algorithm is proposed to optimize the path of solder joint detection of circuit board.Throughout the research status at our countryand abroad, path planning technology has made great progress, but in the complex environment, how to successfully model, how to plan the path is still a very important problem.In addition, the research on virtual assembly path planning is less successful, and the practical application is even less.

In this paper, an improved deep reinforcement learning algorithm is proposed to solve the problem of narrow space in complex environment and through the experiment to compare andverify.

2. Algorithm Model

In this paper, the deep Q-network algorithm in deep reinforcement learning is used to solve the path planning problem of virtual assembly, and the fuzzy Bayesian decision-making algorithm is proposed to solve the problem of exploration and utilization.Firstly, the prior knowledge in exploration is obtained by using fuzzy comprehensive evaluation method to synthesize various factors.Then the Bayesian decision algorithm updates the posterior distribution based on the known prior probability distribution.Finally, actions are selected according to posterior distribution to generate decisions to optimize the path planning of virtual assembly.

2.1. Deep Q Network Algorithm

Deep Q network algorithm [13] is mainly composed of Q-learning and deep convolutional neural network (CNN) [10,11,12]. Q-learning is a reinforcement learning algorithm with different strategies.

The learning process of Q-learning is as follows:

In the first stage: the agent chooses action a to interact with the environment according to the

□ -greedy algorithm under the condition of randomly given state s;

In the second stage: after the agent performs action a, the current state changes and enters the next state. At the same time, the environment will give the agent an immediate feedback (reward or punishment).

In the third stage, the R matrix is constructed with the state as the row and the action as the column, and the real-time feedback of the environment to the agent after the action a is executed in the state s is stored in the R matrix;

The fourth stage: according to R matrix, P matrix is calculated to guide the agent's action. The P matrix is composed of state and action key value pairs (state action value function, Q value).The greedy strategy is used to evaluate and improve the Q value, and the formula (1) is used to update the Q value and converge to the optimal Q value;

The fifth stage: according to the P matrix, select the action with the largest value function (Q *, the best return) in each state, and generate the action sequence.The update formula of state action value function is as follows:

$$Q(s_t,a) \leftarrow Q(s_t,a) + \alpha[R(s_t,s_{t+1},a) + \gamma \max_{a \in A} Q(s_{t+1},a) - Q(s_t,a)] \quad (1)$$

In Q-learning, a Q-function Q (s, a) is defined to represent the feedback that can be obtained by taking action a in state S. the Q-value is constantly updated by iteration. If the Q-function is accurate enough and the environment is certain, the strategy of selecting the maximum Q-valueaction can be adopted.

The traditional action value P matrix is mainly realized by the lookup table [14], and the Q value is stored in a Q table.However, when the state increases, the number of agents' actions increases exponentially, which makes the data difficult to store.

2.2. Fuzzy Bayesian Decision Algorithm with Prior Knowledge

Quantitative evaluation of various factors to make decisions.Fuzzy Bayesian decision algorithmwith prior knowledge [15,16,17],The steps are as follows:

Step 1: construct the factor set U of fuzzy comprehensive decision-making method, and the factor set u is the set of all factors affecting the decision-making. Set the action set a and state set s of Q learning, and divide the state set s according to factor set U;

Step 2: construct fuzzy evaluation matrix Ef and weight set W according to experts' experience.According to the factor set U and fuzzy evaluation matrix Ef, through the weight set W, all States si ,i=1,2,…,k. The superiority degree of each decision under K is given by the superiority vector Bi, i=1,2,…,, K means.Then, the Bi of each state si is normalized, and the result is used as the prior knowledge of Q learning to initialize the Q value of state si;

Step 3: start Q learning. In the time step t, according to the current state st, select action aj to reach the new state st+1, get an immediate return $r(s_t, a_j)$, and update the Q value;

In the learning process, Boltzmann method is used to calculate the probability $p(a_j)$ to select the action randomly. Probability $p(a_j)$ is defined as follows:

$$p(a_j)=\frac{\exp[\frac{Q(s,a_j)}{T}]}{\sum_k \exp[\frac{Q(s,a_k)}{T}]} \tag{2}$$

Where $Q(s,\ a_j)$ is the Q value of the state s-action aj pair, T is the temperature parameter in the annealing process. The larger T is, the greater the probability of random sampling is;

Step 4: the posterior probability distribution is calculated by Bayes formula, and the improvedBayes formula (5) is obtained by combining the action and state in Q-learning with formula (4),The definition is as follows:

$$\theta_{t+1}=\theta_t+\alpha\left[r+\gamma\max_{a'}Q(s',a';\theta)-Q(s,a;\theta)\right]\nabla Q(s,a;\theta)$$

$$p(a_j\mid s_t)=\frac{p(s_t\mid a_j)\cdot p(a_j)}{p(s_t)} \tag{3}$$

3. Analysis of Experimental Results

The experiment is based on OpenAI Gym environment library and implemented by TensorFlow, a simulation environment with unreachable state is established randomly based on OpenAI Gym environment library.

The experiment shows that although the fuzzy Bayesian deep Q network algorithm takes more time to explore the environment and consumes a certain amount of time to explore the environment, the path will be re optimized immediately. When the number of steps reaches 120, the path tends to be stable.

In order to verify the planning efficiency of deep q-network in virtual assembly path planning in narrow space, this paper compares the experimental results of the traditional fast expanding random tree algorithm(Rapid-exploration Random Tree, RRT) [18,19,20] and the improved fuzzy Bayesian depth q-network in the same simulation environment.

Table 1. Comparison of planning efficiency between RRT algorithm and fuzzy Bayesian deep Q network algorithm

	50%	40%	30%	20%	10%
RRT[25]	40.0184527	39.4535781	34.5006577	31.5828547	321.0748215
RRT[100]	3.1327235	3.0223597	3.2958478	3.9965231	30.6325810

FB-DQN	2.6912584	2.7098782	2.8379538	2.6258714	2.5997448

The above table shows the efficiency comparison chart of the two algorithms. It can be seen from the comparison that the planning efficiency of RRT algorithm depends on the setting of search step size. However, in this experiment, when the step size is too large, it will deviate from the minimum width of the free space in the global map, resulting in planning failure. Although the path planning time is reduced to a certain extent, the success rate also decreases. However, the fuzzy Bayesian depth Q-network algorithm does not need to set the search step size, and has stable planning efficiency and short planning time under different global traffic degrees. In the narrow space, using the fuzzy Bayesian depth Q network algorithm for path planning has better planning efficiency. Therefore, in the narrow space, using deep Q-network algorithm to solve the path planning problem of virtual assembly has better trafficability and planning efficiency.

4. Epilogue

In this paper, fuzzy Bayesian decision-making algorithm and fuzzy Bayesian depth Q-network algorithm with prior knowledge are proposed. The experimental results show that the proposed algorithm has good trafficability and planning efficiency for virtual assembly path planning, and achieves the expected goal.

References

[1] Singh Y, Sharma S, Sutton R, et al. Optimal Path Planning of an Unmanned Surface Vehicle in a Real-Time Marine Environment using a Dijkstra Algorithm[C]// The International Conference on Marine Navigation and Safety of Sea Transportation. 2018:399-402.

[2] Fu B, Chen L, Zhou Y, et al. An improved A* algorithm for the industrial robot path planning with high success rate and short length[J]. Robotics & Autonomous Systems, 2018.

[3] Altinoz O T, Yanar T A, Ozguven C, et al. Improved non-Probabilistic Roadmap method for determination of shortest nautical navigation path[C]// International Conference on Electrical and Electronic Engineering. IEEE, 2017:261-266.

[4] Zeng N, Zhang H, Chen Y, et al. Path planning for intelligent robot based on switching local evolutionary PSO algorithm[J]. Assembly Automation, 2016, 36(2):120-126.

[5] Pérezhigueras N, Caballero F, Merino L. Learning Human-Aware Path Planning with Fully Convolutional Networks[C]// IEEE International Conference on Robotics and Automation. IEEE, 2018.

[6] Flavigne D, Taix M, Ferre E. Interactive motion planning for assembly tasks[C]// The, IEEE International Symposium on Robot and Human Interactive Journal of agricultural machinery Communication, 2009. Ro-Man. IEEE, 2009:430-435.

[7] Liu Jianhua, Yang Jianguo, Liu Huaping, Geng Peng, Gao Meng.Global path planning method for mobile robot based on potential field ant colony algorithm[J].Journal of agricultural machinery,2015,46(09):18-27.

[8] Lillicrap T P , Hunt J J , Pritzel A , et al. Continuous control with deep reinforcement learning[J]. Computer Science, 2015, 8(6):A187.

[9] Wang Z, Schaul T, Hessel M, et al. Dueling network architectures for deep reinforcement learning[J]. 2015.

[10] Elbaz G, Avraham T, Fischer A. 3D Point Cloud Registration for Localization Using a Deep Neural Network Auto-Encoder[C]// Computer Vision & Pattern Recognition. 2017.

[11] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016:770-778.

[12] Llinas J.,Hall D.L..An introduction to multisensor data fusion[J].1997,85(1).

[13] Pan Y , Zhang L , Li Z W , et al. Improved Fuzzy Bayesian Network-Based Risk Analysis With Interval-Valued Fuzzy Sets and D-S Evidence Theory[J]. IEEE Transactions on Fuzzy Systems, 2019, PP(99):1-1.

[14] RODRIGUEZ S,TANG X,LIEN J M,et al.An obstacle- based rapidly- exploring random tree:IEEE

International Conference on Robotics and Automation[C].

[15] Zhuxia, Chenrenwen, Xudongxia,etc. Solder joint inspection path planning method based onImproved Particle Swarm Optimization[J]. Journal of instrumentation, 2014(11):2484-2493.

[16] Shuokaiping, Zhangxiaoshun, Xutao,etc. Joint optimal scheduling of multi energy systems based on knowledge transfer Q-learning algorithm[J]. Power system automation, 2017, 41(15):18-25.

[17] Wang Xuesong, Zhu meiqiang, Cheng Yuhu. Reinforcement learning principle and its application [M]. 2014.

[18] Cheng Xiaoping. Fuzzy Bayesian networks based on α - cut sets [J]. Computer science, 1999, 000 (006): 65-66,90.

[19] Liu Xiaoqian, Zhang Hui, Wang Yingjian. Path planning algorithm based on improved RRT [J]. Automation technology and application, 2019 (5): 96-100.

[20] Liu Chengju, Han Junqiang, Ankang. Dynamic path planning of roboc cub robot based on improved RRT algorithm [J]. Robot, 2017 (39): 8-15.