

Research on Learning from Demonstration of Mobile Robot with Autonomous Navigation

LI Yan-fei^{1,a} and ZHANG Wen-zhi²

^{1,2}*School of Mechanical Engineering, Inner Mongolia University of Technology, Hohhot 010051, China*

Abstract: A new method of batch learning from Demonstration is presented, in order to solve the problem of mobile robot independent navigation. According to the actual situation, the model of the Learning from Demonstration is given, and the neural network is used to realize the robot's learning. Considering that the single artificial neural network cause dimension disaster, we designed the Learning from Demonstration model which is the coexistence of multi-neural network and is dynamic switching. The simulation results demonstrate that mobile robot independent navigation is realized.

1 Introduction

Mobile robot should be able to perceive changes in the surrounding environment and in accordance with changes in the environment appropriate to adjust their action path and behavioral strategies [1]. In the field of military, mobile robot technology has been applied to a variety of advanced unmanned early warning aircraft, demining robots; In the civil field, domestic mobile, entertainment, medical and other types of mobile robots more and more people in the field of vision. In short, the mobile robot has a very broad space for development and application prospects. However, navigation is a necessary problem to be solved by the mobile robot, which determines the action set of the mobile robot from the initial point to the target point, and avoids the collision with the obstacle [2,3]. The existing algorithms include grid method, potential force method and fuzzy control method. These algorithms must be designed by the professionals according to the surrounding environment of the robot, and the environment changes will affect the navigation and obstacle avoidance of the mobile robot. And even the need to rewrite the control procedures by experts, bringing expensive human and material resources [4,5]. Aiming at the existing navigation algorithms of mobile robots, A navigation controller for mobile robot based on batch demonstration learning is proposed. According to the frame of demonstration and the actual situation of the mobile robot, a mobile robot model based on demonstration learning is designed. And the neural network learning algorithm is used to compensate the non-linear term between the environment state and the action in the model. Using the control method proposed in this paper, a two-wheeled mobile robot is used to simulate an arbitrary path in an obstacle-free environment in order to realize autonomous navigation.

2 Demonstrate learning model of Mobile robot

2.1 Frame of Batch of learning from demonstration

In batch learning, all presenter sample data are collected prior to learning, and the learning update itself often uses the mathematical properties of the strategy evaluation value M. The batch learning process is shown in Fig1. From human brain

^a Corresponding author: 839765543@qq.com

strategies to collect large amounts of human state space data [6]. The human state space is mapped into a common task space by the human brain's task space operator. Through the theoretical analysis of the strategy evaluation value M , the effective data in the general state space is selected and imported into the updating operator U , and finally the robot control strategy is derived through the robot task space operator.

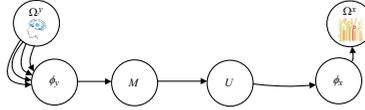


Figure 1. Batch learning

2.2 Overall Model of Mobile Robot Based on Batch Demonstration Learning

The essential problem of mobile robot based on batch demo learning is to solve the nonlinear mapping between environment states and actions. Combining the frame structure of batch demonstration learning and the actual situation of mobile robot, the whole model of mobile robot based on batch demonstration learning is designed. As shown in Figure 2.

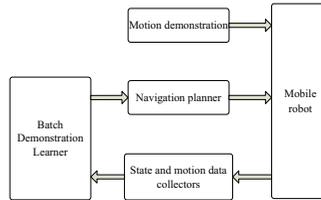


Figure 2. Demonstrates learning the mobile robot model

2.3 Design of Navigation Planner

In the overall model of the mobile robot, the navigation planner is the most critical module for the robot to realize self-navigation. Its role is to achieve the robot state of the environment and the implementation of the nonlinear mapping between the actions. There is no fixed pattern for this many-to-one mapping, so it is almost impossible to find the formulas between them. At present, artificial neural network (ANN) has been widely used in the development of nonlinear models. It is especially suitable for applications where input and output are not well defined. It is feasible to apply it to the navigation planner in Fig2. Considering the complexity of the navigation planning function, if a single neural network is used to complete the function of the navigation planner, the artificial neural network will be too large, the large n_e . In view of this feature, the navigation planner is divided into several smaller planners, and a classifier is added to select the corresponding subordinate planner to control the robot sailing movement according to the state of the robot. According to this idea, the navigation planner can be designed as shown in Figure 3 the overall structure.

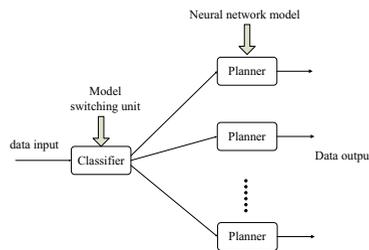


Figure 3. Learning model overall framework

In the above diagram, each planner is implemented with a small-scale neural network to form the structure of the multi-neural network. Each neural network can use the same structure, but the neural network training using different data sets, of course, the network structure can also be used in different forms of structure. Although each neural network is not perfect, it can only generalize some types of robot environment states and motion maps, but through the model switching unit constructed before the mulch-neural network, the dynamics of each neural network model in the robot running process conversion. Dynamic conversion to make the performance of each neural network perfect play to achieve the proper function of the navigation planner. In order to determine the number of planners and planner functions in Figure 3, it is necessary to analyze the state information obtained by sensors on the actual robot, the characteristics of navigation target points and so on. The robot used to test the learning effect of the mobile robot is a two-wheeled robot equipped with three distance detection sensors with an angular spacing of 20° between the three sensors, which are located in front of the robot, front left and front right respectively. The state of the machine can be divided into eight states according to whether the three sensors detect an obstacle: NNN, NNE, NEN, ENE, NEE, ENE, EEN, and EEE, where E is the detected state Existing, N indicates that no obstacle has been detected (Nothing). In accordance with the overall structure of the design idea, each small planner uses an artificial neural network to replace. The entire navigation planner consists of eight neural networks and a classification switching unit. And represent the obstacle distance values detected by the three sensors on the mobile robot, respectively, and is the steering angle of the mobile robot, which can be used as the input signal of the eight neural networks. Thus, the navigation planner shown in Figure 3 can be further refined to form a detailed plan of the navigation planner as shown in Figure 4. The module switching unit dynamically triggers one of the eight neural networks according to the output value of the three sensors, and outputs the control parameter for controlling the steering of the robot.

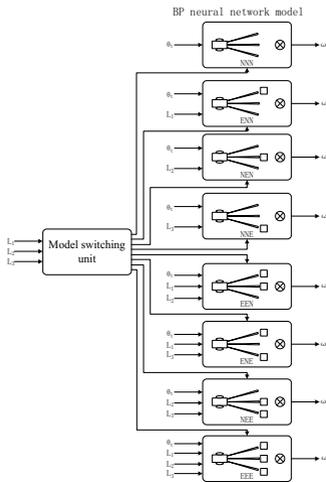


Figure 4. Detailed block diagram of the navigation planner

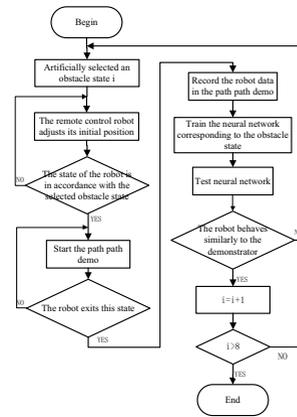


Figure 5. Demonstrate learning flow chart

In order to make the mobile robot perform the task of demonstration learning, the weights of the internal nodes of each neural network in Figure 4 must be updated by using the data extracted by the state and action data collectors in Figure 3 after the demonstrator is finished. Can be used BP algorithm, the field of intelligent control is widely used in a kind of neural algorithm [7]. It can store and generalize this complex input-output mapping relationship and control the output precision of the network by training the steepest descent learning rule under the condition that the complex input-output mapping relationship is difficult to be expressed by mathematical function [8, 9].

3 Simulation experiment

3.1 Demonstrate learning process

In the MATLAB simulation process, the virtual mobile robot sensor distribution as shown in Figure 5, mobile robot, obstacle and target sphere position of any of the three placed. The learning flow is shown below.

3.2 Analysis of learning result

With the human hands demonstrate the behavior of the ongoing remote control, eight kinds of obstacle status will produce a large number of presentation data. Using this data, the corresponding neural network is trained repeatedly, and finally eight neural network models are obtained which accord with the demonstrator behavior. Figure 6 shows 2D neural network control graphs without obstacle state, 3D neural network control diagram with single obstacle and double obstacle.

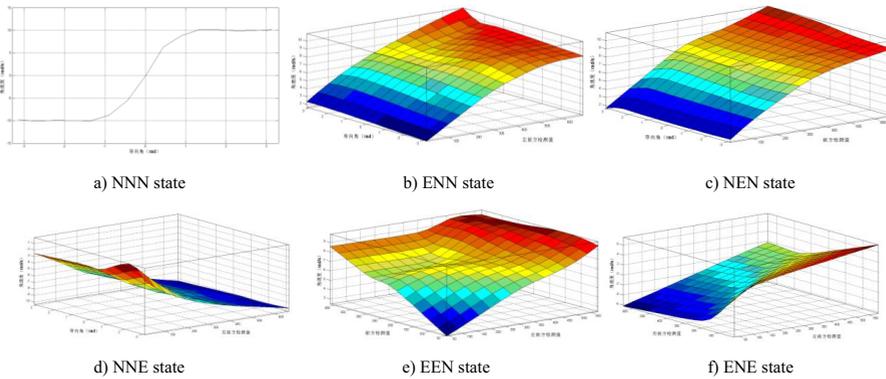


Figure 6. Curve and surface diagram of neural network control

The data sources in each image in Figure 6 are extracted from the trained neural network controller, which can reflect the performance of the demonstration learning controller. Some typical data points are selected from the figure to illustrate:

As can be seen from the figure a, the guide angle is -3.14 and the angular velocity output is -10; the guide angle is 3.14 and the angular velocity output is 10; It is shown that the NNN neural network can produce a large angular velocity when the moving direction of the mobile robot and the deviation angle of the target sphere are large.

From Figure b, c can be seen in the case of a single obstacle, the sensor detects the value of 700, the guide angle of any given angular velocity output in the vicinity of 10; d is similar, it is clear that the closer the mobile robot is to the obstacle, the larger the angular velocity of the single obstacle neural network.

Can be seen from the e graph, the left front detection value and the front detection value as long as there is a 400, the angular velocity output in the vicinity of 10; It is shown that the angular velocity of the EEN neural network increases as the mobile robot approaches the obstacle.

As can be seen from the f graph, the left front detection value and the right front detection value are 500, the angular velocity output in the vicinity of 0. It is shown that the ENE neural network can control the mobile robot to move along the obstacle.

3.3 Simulation and experimental platform testing

In order to test the performance of the self-navigation controller, the obstacle in 3D virtual scene of the simulation platform is rearranged and two kinds of complicated test environments are designed. The simulation results are shown in Figure 7.



Figure 7. Virtual Robots in a Static Environment and in a G - Type Environment

After testing the performance of the demonstration learning control model, the subsystem is built on the platform of the two-wheeled mobile robot, and the obstacle environment is arranged. The demonstration experiment is completed on the experimental platform. As shown in Figure 8, the robot successfully avoids a plurality of obstacles and reaches a predetermined target point.

Experimental results on the experimental platform show that the self-navigation control strategy of mobile robot based on batch demonstration learning has better navigation performance.

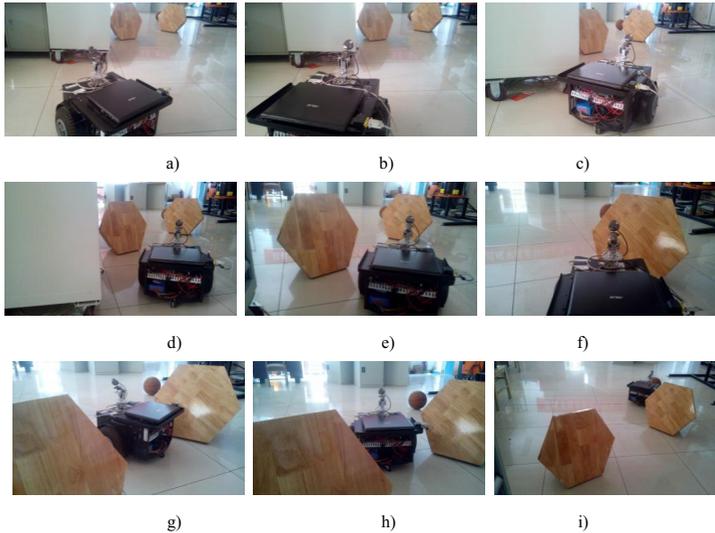


Figure 8. Mobile robot test platform path navigation test

4 Conclusions

Aiming at the characteristics of mobile robot demonstration learning, artificial neural network is proposed to realize the learning action of mobile robot to demonstrator. Considering single artificial neural network to realize complex navigation demonstration learning will make neural network too complex or cause neural network convergence Neural network, and each neural network is relatively simple, only to achieve a certain state of the mobile robot learning action, the robot in its mobile In the process, according to their work status changes at any time to switch to the appropriate neural network, at any time there is only one neural network in working condition. Experiments show that the neural network with this structure has a faster convergence rate. The learning model is simulated by the simulation platform, and the test result is good. Finally, the test results on the real robot show that the self - navigation control method based on batch demonstration learning is feasible.

References

1. C.Merikli, Manuela Veloso, H. Levent Akin. A. IEEE-RAS,(2010).
2. H.Bener Suay, Sonia Chernova. A comparison of two algorithms for robot learning from demonstration[C]. Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics. IEEE Press,(2011).
3. C.Sun, Wei He, Weiliang Ge, Cheng Chang. Adaptive Neural Network Control of Biped Robots[C]. Man and Cybernetics: Systems. IEEE Press,(2016).
4. H.Niu, Niu Wang, Nan Li. The adaptive control based on BP neural network identification for two-wheeled robot [C]. World congress on intelligent control and automation. IEEE Press,(2016).

5. Santiago Morante, Juan G. Victores, Carlos Balaguer. Automatic demonstration and feature selection for robotlearning [C] .Humanoid Robot. IEEE-RAS,(2015).
6. Sonia Chernova, Manuela Veloso. Multi-thresholded approach to demonstration selection for interactive robot learning [C] . Proceedings of the 3rd ACM/IEEE International Conference on Human-Robot Interaction. IEEE Press,(2008) .
7. S.Jia, Quan Qiu, Junmin Li; You Li, Yue Cong. BP neural network based localization for a front-wheel drive and differential steering mobile robot [C] . Information and Automation. IEEE,(2015).
8. Maria Koskinopoulou, Stylianos Piperakis, Panos Trahanias. Learning from Demonstration facilitates Human-Robot Collaborative task execution [C] . Human-Robot Interaction.ACM/IEEE,(2016).
9. Hiroaki Arie, Takafumi Arakaki, Shigeki Sugano,et al. Imitating others by composition of primitive actions: A neuro-dynamic model [J] . Robotics and Autonomous Systems,(2012).