Applied Mathematics & Information Sciences
An International Journal

A Navigation Method based on BA-POMDP Algorithm

Yong Tao
Beihang University, Beijing 100191 P. R. China
Email: taoyong@me.buaa.edu.cn

Received: 09 Mar. 2012; Revised 25 Jul. 2012; Accepted 04 Aug. 2012

Abstract: This paper presents a robot navigation method based on hierarchical POMDP and Bayesian Algorithm (BA-POMDP Algorithm) in uncertain environments. A successful and efficient robot navigation method in dynamic environments requires predicting that the uncertainties of state of events as well as obstacles. A novel procedure accounting for both state transition and observation uncertainty in the navigation process is presented. In order to solve the problem in dynamic planning programming that is associated with robot navigation in uncertain environments, we present BA-POMDP algorithm that integrate prediction, estimation and planning while also properly fuse weights by mapping between fusion weights and the immediate environmental configuration. The algorithm is implemented and tested onboard that the elderly companion robot achieves autonomous navigation. Experiments from dynamic scenarios illustrate the effectiveness of the BA-POMDP algorithm.

Keywords: Navigation method, hierarchical POMDP, BA-POMDP model, elderly companion robot.

1 Introduction

With the aging of the society, the number of elderly is increasing. Limited mobility and high dependence on others affects people's self-confidence negatively. In accordance to current society trends and the claim for intelligent assistive devices, many smart systems have been developed in the last decades, for instance: robotic walkers [1], smart blind sticks [2] and several wheelchair platforms.

Robots designed for the home are a growing industry from both a research and commercial perspective. Robot companions are expected to communicate with non-experts in a natural and intuitive way. A companion robot will generally be acting in an unstructured environment, such as a private home or a rest home for old people, with people roaming around. Companion robots in the home should ideally be able to perform a wide array of tasks including educational functions, home security, diary duties, entertainment, message delivery services and so on [3-6]. Currently, there are no robots that are able to perform a combination of these tasks efficiently, accurately and robustly. Since it is not desirable to rely on pervasive sensor technology distributed throughout the environment, the companion robot has to carry all sensing devices on board. At the same time, there has been little research to date in terms of assisting elderly people with cognitive tasks, such as remembering medication schedules.

Partially Observable Decision Processes (POMDP) are techniques for calculating optimal control actions under uncertainty [7-8], and in mobile robotics in particular. They are useful for a wide range of real-world domains where joint planning and tracking is necessary, and have been successfully applied to problems of robot navigation and robot interaction. However, standard methods for planning, inference, and learning with POMDP all take time exponential in the number of states, making them impractical for-large problems.

This paper studies a robotic navigation method based on BA-POMDP algorithm in a dynamic changing environment. The model can account for both state transition and observation uncertainty. The behaviors for autonomous navigation in an indoor environment are given: goal seeking behavior, wall following behavior, and obstacle avoidance behavior. Each behavior is designed by using hierarchical POMDP approach to achieve respective navigation task. The output of BA-POMDP of each behavior reacts to environmental information from ultrasonic and laser sensors. The proper fusion weights are given by mapping between fusion weights and the immediate environmental configuration. The method is implemented and tested onboard a companion robot to achieve autonomous navigation. The experimental results verify the effectiveness of the navigation method.

2 Introduction of the Elderly Companion Robot

The robot is built on top of a Nomad Scout differential drive mobile base, equipped with 8 ultrasonic sensors and 8 infrared sensors. It is also equipped with two on-board PCs, connected to the Internet via a 2Mbit/sec wireless Ethernet link. A bright, touch-sensitive color display is mounted conveniently at approximate eye height for sitting people. The current prototype companion robot is shown in Figure 2.1.

Figure 2.1: Side view of the companion robot

There is a color CCD camera with an approximate aperture angle of 100 degrees. It is also equipped with a speaker system and a microphone array, necessary for recording and synthesizing speech. The battery lifetime is approximately 60 minutes.

3 Hierarchical POMDP

The basic idea of the hierarchical POMDP is to decompose the action space into smaller chunks. Since the state is not fully observable, the hierarchical POMDP is based on action hierarchy. It's assumed that the maps are already available and that the topological map representation is hand coded and programmed. Figure 3.1 illustrates the basic concept of the action hierarchy.

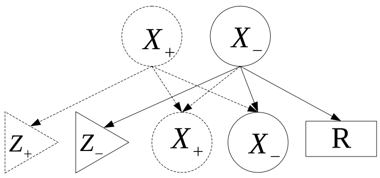

Figure 3.1: Dynamic belief network

Formally, an action hierarchy is a tree, where each leaf is labeled by an action from the action set of POMDP. Each action $a \in A_0$ must be attached to at least one leaf, A_0 is the primitive action. In each internal node, there is an abstract action [9].

The hierarchical POMDP model is to use the hierarchical action set to replace the full POMDP. Through the hierarchical action set, we achieve a collection of POMDPs that are smaller than the original POMDP, yet collectively define a complete policy.

Given that the action hierarchy spans a collection of separate POMDP subtasks, we can independently optimize an independent policy for each subtask, such that we obtain a collection of corresponding local policies.

With a subtask P_i, action set A_i, state set S_i and observation set O_i, where the state space is corresponding to state features $X_i = \{X_1, X_2, ..., X_m\}$ and the observation space is corresponding to features $Z_i = \{Z_1, Z_2, ..., Z_n\}$. We now consider a case where the state features can be divided into two different sets, X_+ and X_-, and similarly observation features can be divided into two disjoint sets, s and Z_-. The state features s and observation features Z_+ are irrelevant to subtask P_i if $\forall a_j \in A_i$:

$$R(X_+, X_-, a_j) = R(X_-, a_j) \tag{11}$$

In hierarchical POMDP model, abstraction can be applied independently to each subtask - thus reducing $|S|$ and $|O|$ - without influencing the policy optimization any further than what is attributed to the action decomposition [10].

4 Navigation Method based BA-POMDP Algorithm

The standard Bayesian reinforcement learning framework uses distributions to represent the prior and posterior distributions over the unknown transition probabilities defining the model [11-13]. Distributions are convenient because they represent the probability that some random variable follows a particular discrete distribution, given the number of times each event has been observed thus far. Parameters can be estimated exactly by simply counting the number of times each state transition occurred.

Planning is achieved by specifying an extended MDP model, where the distribution parameters are included in the state space. The transition function models how these parameters are updated given a particular state transition.

When modeling a robot domain as a finite POMDP, distributions can also be used to represent the prior distribution over the unknown transition and observation probabilities. However, an added complication arises due to the fact that the state is not observable, therefore it is not possible to know the exact values of the parameters.

A bayesian model for hierarchical POMDP is proposed by making three changes: we add action nodes U_t , orientation nodes θ_t and allow exit nodes E_t to be multi-valued.

The action nodes represent the movement made by the companion robot. The orientation nodes are present because we now factor X_t^1 into concrete location L_t^1 and orientation θ_t instead of having to duplicate each location four times. We will denote the abstract location X_t^2 by L_t^2. The exit node E_t can take on five possible values, representing no-exit, north-exit, east-exit, south-exit, and westexit. If E_t = no-exit, then we make a horizontal transition at the concrete level, but the abstract state is required to remain the same.

The companion robot navigation systems have the capacity to handle uncertain and imprecise information using the BA-POMDP rules. When a robot navigates in an unstructured environment, it needs to detect its surrounding environment and interprets the environmental information. From sensory information, the robot can obtain the knowledge of its position and the distance to surrounding obstacles. POMDP has been widely used for mobile robot navigation, mainly due to it is capable of offering inference using environmental data, even under motion and sensor uncertainties. There are three behaviors by using BA-POMDP. The inputs of each BA-POMDP are environmental information measured by ultrasonic, visual and laser sensors, and the output is the motion speed between left and right wheels.

There are three navigation behaviors: goal seeking behavior, wall following behavior, and obstacle avoidance behavior. Each behavior takes about 25s to run. During the navigation, the robot tries to provide the information and assist navigation of the category to which the behavior belongs.

A total of 8 subjects were used in this study. All subjects were students or colleagues of our institute. They were not familiar with the use of the robot, and had no prior experience of interaction with the robot.

We exposed 8 subjects to the behaviors, and let them choose what kind of navigation they were comfortable with for each of these. The robot did not give information or navigation service in the pre-study. There are variations of preferences for each subject within the same class.

The robot was initially placed in the middle area of indoor environment, and the subject was asked to sit in front of the robot and interact with it in a natural way. The robot randomly selects one of the six interaction behaviors. After one behavior finished, it randomly selects the next one. The navigation test lasts for 20 minutes.

During the navigation, the H-POMDP behavior adaptation method was running on the companion robot. For the duration of each interaction behavior, or the test of a policy, it keeps the interaction observation and other parameters according to π_{Pi}. The reward function R was calculated for each executed interactive action of the robot using the command input and user head position for the duration of the behavior [14].

The duration starts from just after the robot selects the behavior and it ends at the end of the behavior. A total of ten different parameter combinations were tried before optional policy π_i was calculated and the parameter values updated. The subject did not notice the update of the policy during interaction.

5 Experiment Results

For most of the subjects, at least part of the parameters reach reasonable convergence to stated preferences within 10 min, or approximately ten iterations of the behavior adaptation algorithm. Using the hierarchical POMDP, the high-level decision making of the robot is more effective, and the optimal policy could be computed off-line. The main controller of the companion robot just deals with the states and the corresponding action control.

In hierarchical POMDP, the companion robot uses the information perception actions to predict and confirm the user's intention. In the uncertain environment, we tested the robot navigation method in three separate experiments, each lasting half-day.

Figure. 5.1. Experiments of the robot in a living room

The sequence of images illustrates the major stages of a successful service: from moving throgh the corridor, navigating through the door, avoiding the obstacle, and providing information after successfully reaching the destination. In all navigation experiments, the tasks were performed to completion.

The robot autonomously led 15 full transportation services for 8 different testers. The goal seeking behavior, wall following behavior, and obstacle avoidance behavior, converged to the values near to selected values for 6, 7 and 8 of 8 subjects, respectively.

The experimental results reveal the navigation method can determine the proper fusion weight and response to the dynamic changing environment efficiently. However, there are some mistakes at one of the end of the path. It is caused by the position estimation error of the companion robot. In the future, the more accurate localization design will be done to the robot and the behavior fusion method for more complex tasks will be investigated.

6 Conclusion

This paper presents a robotic navigation method based on BA-POMDP algorithm in a dynamic and changing environment. The model can account for both state transition and observation uncertainty. The proper fusion weights are given by mapping between fusion weights and the immediate environmental configuration. The method is implemented and tested onboard that a companion robot achieves autonomous navigation in a dynamic and changing environment. The experimental results verify the effectiveness of the navigation method.

Acknowledgements

This research is supported by the 863 Program of China under grant No 2008AA040201, also supported by the Fundamental Research Funds for the Central Universities.

References

[1] Majd Alwan, Prabhu Jude Rajendran, Alexandre Ledoux,etc. *Stability margin monitoring in steering-controlled intelligent walkers for the elderly*, in Proceedings of AAAI Fall Symposium (2005).

[2] Kang S, Ho Y, and Hyuk Moon I. *Development of an illigent guide-stick for the blind*, in ICRA, Seoul, Korea, 3208-3213 (2001)

[3] Mitsunaga N., Smith C., Kanda T., etc. *Adapting Robot Behavior for Human-Robot Interaction.* I*EEE Transactions on Robotics*, 24, 911- 916 (2008).

[4] Ocana M., Bergasa L.M., Sotelo M.A., etc. *Indoor robot navigation using a POMDP based on WiFi and ultrasound observations,* in IROS 2592-2597 (2005)

[5] Demeester E., Huntemann A., Vanhooydonck D., etc. *Bayesian estimation of wheelchair driver intents: Modeling intents as geometric paths tracked by the driver*, in IROS.5775–5780 (2006)

[6] Mandel C, Huebner K, and Vierhuff T. *Towards an autonomous wheelchair: Cognitive aspects in service robotics*, in Proceedings of Towards Autonomous Robotic Systems , 165-172 (2005).

[7] Kaelbling L. P., Littman M. L. and Cassandra A. R. *Planning and acting in partially observable stochastic domains, Artificial Intelligence*, 101, 99-134 (1998)

[8] E.Lopez, R.Barea, L.M.Bergasa and M.S. Escudero. *A human-robot cooperative learning system for easy installation of assistant robots in new working environments*, Journal of Intelligent and Robotic System 40, 233-265 (2004).

[9] Martha E. P., Sandra E. and Sebastian T. Pearl: *A Mobile Robotic Assistant for the Elderly*, in AAAI Technical Report WS-02-02, (2002).

[10] Pineau, J., and Thrun, S. *High-level robot behavior control using pomdps*. In AAAI-02 Workshop on Cognitive Robotics, (2002).

[11] R. Dearden, N. Friedman, and N. Andre. *Model based Bayesian exploration*, in UAI, (1999).

[12] M. Duff. Optimal learning: *Computational procedure for*

bayes adaptive markov decision processes. Ph.D. dissertation, University of Massachusetts, Amherst, USA, (2002).

[13] P. Poupart, N. Vlassis, J. Hoey, etc. *An analytic solution to discrete bayesian reinforcement learning*. in Proc. ICML, (2006).

[14] T. Taha, J. Valls Mir´o and G. Dissanayake. *POMDP-based long-term user intention prediction for wheelchair navigation*. in ICRA, 3920-3925（2008）.

Yong Tao received the MS degree from Chongqing University in 2005, and the PhD degree in Mechanical and electronic engineering from Beihang University in 2009. His research interests are in the areas of robot automation, mechatronics and embedded systems.