Available online at www.sciencedirect.com

ScienceDirect

IFAC-PapersOnLine 49-15 (2016) 114–119

FLAP for CAOS: Forward-Looking Active Perception for Clutter-Aware Object Search [1]

Thorsten Gedicke, Martin Günther and Joachim Hertzberg

{tgedicke, martin.guenther, joachim.hertzberg}@uos.de

Abstract: In this paper, we present a system for autonomous object search and exploration in cluttered environments. The system shortens the average time needed to complete search tasks by continually planning multiple perception actions ahead of time using probabilistic prior knowledge. Useful sensing actions are found using a frontier-based view sampling technique in a continuously built 3D map. We demonstrate the system on real hardware, investigate the planner's performance in three experiments in simulation, and show that our approach achieves shorter overall run times of search tasks compared to a greedy strategy.

Keywords: autonomous mobile robots, heuristic searches, knowledge acquisition, path planning, probabilistic models, robot vision, sampling actions, searching systems

1. INTRODUCTION

Mobile service robots need to operate in dynamic environments that are subject to change not only as a result of the robot's own actions but often due to independent change. The robot's knowledge about the physical state of such an environment may get wrong and incomplete. As an example of such a domain, consider a robot waiter in a restaurant. Since there are other waiters and guests involved, who interact with the environment and introduce changes without announcing these changes to the robot, it can never fully rely on the internally presumed state of the world, especially regarding the whereabouts of task-relevant objects. Instead, it needs to actively incorporate explicit sensing actions into its plans to search for objects or to verify beliefs about the current state of the world. View obstructions due to clutter must be accounted for and sensing actions at a variety of locations may be necessary. Systems that reason about such *active perception* actions to enhance their operational capability and flexibility in changing or unknown environments comprise several interrelated aspects and have been studied in diverse contexts.

While our work focuses on bringing the robot's sensor into configurations desired for data acquisition in a collision-free way, interactive perception actions that move occluding objects out of the way have also been considered in recent literature (Dogar et al., 2014; Gupta et al., 2013; Wong et al., 2013). Collision-free active perception systems typically determine a *next best view*, i.e., a sensor configuration that the robot should assume next to optimize a utility function that incorporates the expected information gain and often also predicted costs, which typically amount to the action's execution time. Next best view algorithms can be seen as a variation of Connolly's (1985) "Planetarium Algorithm", which compares simulated range images to successively select the view pose with the largest area of yet-unknown space in the image. While the objective of exploration is different from object search, it can be argued that a general object search system covers the use case of exploration in the special case that it terminates only after exhausting the search space and is not guided by prior knowledge regarding some target object.

We present a system for object search and exploration in cluttered environments that aims to minimize expected total task execution times using a continual planning method that exploits partial knowledge of the environment and a probabilistic model of the target object's location. The software is implemented using the ROS framework and designed as an active perception module for the artificial cognitive system RACE (Hertzberg et al., 2014). It is available as open source at `https://github.com/uos/uos_active_perception`.

After reviewing related work in Sec. 2, we present our approach in Sec. 3. We show experimental results in Sec. 4 and conclude with a discussion of future work in Sec. 5.

2. RELATED WORK

Establishing a prior hypothesis regarding probable locations of the target object allows to direct the search process towards such locations and enables the system to work in large environments where uninformed search would be infeasible. Ye and Tsotsos (1999) [also Shubina and Tsotsos (2010)] maintain a probability distribution function for a single target object that assigns to each cell in a discretized workspace the degree of belief of the target being located there. Approaches that have been examined to obtain the required world knowledge include semantic probabilistic environment models (Kunze et al., 2012), visual saliency (Rasouli and Tsotsos, 2014), as well as relational models of

[1] All authors are with the Institute of Computer Science, University of Osnabrück, Albrechtstr. 28, 49076 Osnabrück, Germany. Martin Günther and Joachim Hertzberg have a second affiliation with the DFKI Robotics Innovation Center, Osnabrück Branch, 49076 Osnabrück, Germany. This work is supported by the RACE project, grant agreement no. 287752, funded by the EC Seventh Framework Programme theme FP7-ICT-2011-7.

2405-8963 © 2016, IFAC (International Federation of Automatic Control) Hosting by Elsevier Ltd. All rights reserved.
Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2016.07.718

object classes (Wong et al., 2013), affordances (Moldovan and Raedt, 2014), and spatial relations (Kunze et al., 2014; Anand et al., 2011).

Once a region of interest has been identified, the robot needs to select a next best view, i.e., a sensor configuration that brings that region into sight, considering criteria that usually include visibility of unexplored space and sensing costs, e.g., the time needed to travel to the sensing pose. Ye and Tsotsos (1999) and Shubina and Tsotsos (2010) decompose view selection into robot relocation and choice of sensor configuration. Sensor configurations are selected iteratively at each location based on their expected probability to detect the target until travel to a new location becomes necessary. Yamauchi (1997) proposes to guide autonomous exploration in 2D by targeting *frontiers*, i.e., boundaries between free space around the robot and unknown space beyond, a concept which Surmann et al. (2003) extend to autonomous mapping in 3D by sampling potential sensing poses in a 2D projection of 3D space. Blodow et al. (2011) evaluate frontiers in full 3D only for promising view pose candidates previously identified in a 2D projection of the map. Dornhege and Kleiner's (2013) frontier-void-based approach explicitly associates clusters of unknown map cells with neighboring clusters of frontier cells to identify high visibility view poses with six degrees of freedom. The original deterministic form of this approach is modified in later research by Dornhege et al. (2015) to employ sampling. While these systems operate in a cycle of view selection and sensor relocation, Shade and Newman (2011) present an approach that merges view selection and path planning to generate smooth exploration trajectories towards frontiers using a potential field method inspired by fluid dynamics.

The above cited systems, with the exception of Dornhege et al. (2015), select views using a greedy strategy. Computing an optimal search/exploration strategy is NP-hard and related to the set cover, art gallery, and traveling salesman problems (Ye and Tsotsos, 1999; Sarmiento et al., 2003). Planning multiple steps ahead is necessary in certain applications without regard for optimality, e.g., to avoid traversing unknown space to scan a target (Renton et al., 1999) or to expose unknown space behind multiple layers of concealment (Gupta et al., 2013). Several approximate methods to generate more efficient plans have been examined that rely on divide and conquer (Dogar et al., 2014), pruning heuristics (Sarmiento et al., 2003), and a set cover/TSP decomposition of the problem (Dornhege et al., 2015).

Planning systems for object search have been shown in experiments to provide useful results without prohibitive computational cost (Sarmiento et al., 2003; Dornhege et al., 2015). However, the advantage of planning compared to a greedy approach is outweighed by computational cost when a certain amount of replanning is considered (Dornhege et al., 2015). Since search/exploration systems are likely to encounter circumstances that deviate from the presumed world state, it is desirable to lower the computational burden and make continual planning feasible.

Since it does not introduce any restrictions on sensing poses and allows dynamic scaling of view pose density, sampling seems to be a favorable approach to generate sensing pose candidates that has already been used successfully for forward-looking search (Dornhege et al., 2015). The process can be guided towards promising locations using the concept of frontiers (Blodow et al., 2011).

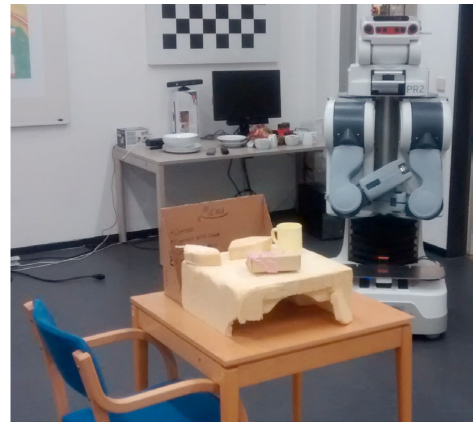

Fig. 1. PR2 during a search process. The target volume on top of the table has several occlusions and a cavity.

This paper contributes three innovations for mobile object search and exploration. First, we present an efficient anytime system for 3D view computation that uses frontier-based sampling to find next best view candidates. Second, we develop an object search strategy that plans multiple sensing actions ahead of time to minimize the expected search time using a predictive model of robot movement speed and a probabilistic model of target object locations. Third, we show the performance of the search planner to be suitable for offline as well as continual planning and demonstrate the complete system running in the real world.

3. APPROACH

Our system searches for an object of a target class (e.g., a mug) within a given target region. It is designed as a subcomponent of a cognitive architecture that is capable of recognizing individual objects when they are detected by the sensor and maintains a belief state regarding probable object locations. When the object search component is triggered by the system, it is given a set of bounding volumes (e.g., a box that encompasses the top of the table shown in Fig. 1) and the degree of belief for each of these volumes to contain an object of the target class. The actions available to the robot are navigating to a new position (using a path planner), pointing the head, and raising or lowering the telescoping torso, and any combination of these may be used to transition between view poses. We rely on the system to terminate the search process when the target is detected within the sensor's field of view.

Fig. 2 shows an overview of the data flow between the involved components and highlights the active perception components that are subject of this paper in a darker shade. A 3D map is continually built from sensor data and is used by a view sampling module, which computes a set of possible view poses along with the expected information gain for each pose. These samples are per request fed to an object search planning and execution module, which aims to minimize the expected time until the target is found and iterates between requesting view samples, planning, moving the robot, and sensing.