Information Sciences 282 (2014) 92–110

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Multi-UAV target search using decentralized gradient-based negotiation with expected observation

Pablo Lanillos [a,*], Seng Keat Gan [c], Eva Besada-Portas [a], Gonzalo Pajares [b], Salah Sukkarieh [c]

[a] Department of System Engineering and Automation, Complutense University, Madrid 28040, Spain
[b] Department of Software Engineering and Artificial Intelligence, Complutense University, Madrid 28040, Spain
[c] Australian Centre for Field Robotics (ACFR), University of Sydney, NSW 2006, Australia

ARTICLE INFO

Article history:
Received 25 October 2012
Received in revised form 11 April 2014
Accepted 26 May 2014
Available online 11 June 2014

Keywords:
Cooperative search
Decentralized decision making
Probabilistic reasoning
Unmaned air vehicle

ABSTRACT

This paper presents a novel approach for the coordination of a team of autonomous sensor platforms searching for lost targets under uncertainty. A real-time receding horizon controller in continuous action space is developed based on a decentralized gradient-based optimization algorithm and by using the expected observation as an estimate of future rewards. The expected observation is a cost-to-go heuristic that estimates the goodness of the states that the platforms could reach. It permits the decision making algorithm to take into account the information on the whole environment, reducing the time needed to detect the target. The heuristic, modeled as a sensor, allows us to develop a new team utility function with low computational cost and high performance. It can be applied to challenging scenarios such as multi-target search with complex and non-uniform target probability distributions. Through simulation and statistical analysis, we show the advantages of using the expected observation heuristic in multi-vehicle coordination for search applications.

1. Introduction

The ability of mobile sensors to search for specific targets of interest in a large scale environment has an increasing importance in many real-life applications. Examples include wildlife monitoring, search and rescue, area patrol, and tactical reconnaissance. Many of these missions are time-critical, which favors the deployment of multiple mobile sensors. A network of mobile sensors has demonstrated the ability to enhance sensing flexibility and achieve the mission objective in a shorter time period. Decentralization among the sensors further provides scalability, modularity, and redundancy to the network. This reduces the vulnerability to central server failure, which further enhances the overall system robustness. In particular, there is an active research community proposing mobile sensor decision making solutions for target search [37,5,12,4,16,27,30,20,13,21,7] and tracking [15,8,17] problems. These solutions intrinsically exploit the action-perception loop [28,10] to design autonomous cognitive agents based on recent computational and decision models of the human brain [1,24]. This paper focuses on decision making (e.g., selecting the actions according to the perception) for target search problems.

* Corresponding author. Tel.: +34 913944740.
E-mail addresses: planillos@fis.ucm.es (P. Lanillos), s.gan@acfr.usyd.edu.au (S.K. Gan), evabes@dacya.ucm.es (E. Besada-Portas), pajares@fdi.ucm.es (G. Pajares), salah@acfr.usyd.edu.au (S. Sukkarieh).

http://dx.doi.org/10.1016/j.ins.2014.05.054
0020-0255/

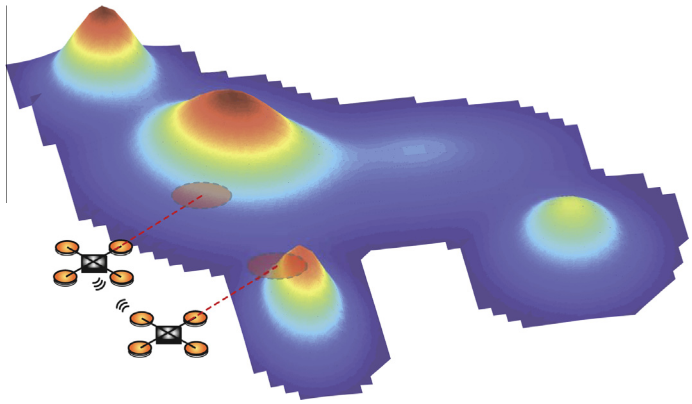

Fig. 1. Target search in a probabilistic scenario using a team of UAVs. The colored height map shows the regions with corresponding probability of finding the targets. This information is used by the team to compute their joint trajectory. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The search problem can be formulated as a probabilistic information gathering task [37,5,16]. Fig. 1 illustrates the problem where a team of two UAVs coordinate their trajectories to detect possible targets. The subjective information about the target locations, shown as a colored height map, is given a priori. It is the information that the team uses, in a probabilistic objective function, to compute their joint action plan. This information or belief[1] is updated every time a UAV makes an observation. It probabilistically describes the target location and helps the team make more precise decisions for finding the target [4].

We can distinguish two kinds of probabilistic measure used as the team objective function: entropy related ones such as the mutual information [15,8,17] or the entropy itself [37,38]; and detection related measures [32], such as the probability of non-detection [9,6,26,14] or the expected time of detection [30,20]. In this paper, we consider the search mission as accomplished once the target is found [9,33]. Entropy based measures usually have difficulties guiding the mobile sensors to find the target in scenarios with general (non-Gaussian) target probability distributions. For instance, in papers such as [16], although they talk about using an entropy measure as the mutual information, in the case of the search task they finally implement the logarithmic probability of detecting the target as the optimizing criterion. Therefore, instead of using entropy related utility functions, we tackle the search problem using a non-detection based objective function. During decision making, we plan sensoring trajectories that minimize the probability of target non-detection, or, equivalently, that maximize the chance of detecting the target.

The complexity of general decision making for the mobile sensor probabilistic search problem is NP-hard [35,3]. That is, solving the optimal sensor trajectory for the length of the mission horizon is intractable in nature due to the exponential growth of the decision tree.[2] Therefore, assumptions are usually made to simplify the problem into a tractable one. Some of these approaches are: greedy approximation [5,37], local optimization [14,27,34], coarse-grained representation of the environment [20] and sampling approximations [21].

Computing the optimal solution for an infinite horizon probabilistic search problem is intractable. However, the search problem can be successfully tackled by simplifying it to a finite horizon problem [5]. It is critical to balance the algorithm's computational time and action myopicity during the action optimization process. It is not a coincidence that although in the related literature the utility functions are designed for N-length action horizons, the first implementations were simplified to 1-step ahead [15,16,6] or 2-step ahead [37,38] controllers due to issues of computational tractability. It is worth highlighting that while the first group [15,16,6] were totally myopic, the second ones [37,38] included a discounted expected reward to reduce the myopicity of the action.

More recent algorithms, such as those proposed in [27,14,20,21], are able to efficiently compute actions N-steps ahead. They can be classified into two groups, according to the representation of the action space. The first group, the discrete action algorithms [20,21], are able to tackle dynamic target scenarios using a sampling-based coarse-grained resolution that trades off between action optimality and computational time. However, discrete optimization is not a real solution, because the decision variables cannot have continuous values, making the controller more complicated to design if the input signals do not have the same granularity. The second group, continuous action algorithms [27,14], are designed to tackle static target scenarios using decentralized optimization.[3] None of the N-step ahead action controllers take into account future

[1] This belief is also known in the literature as the certainty grid [16] or the uncertain probability map [4].

[2] Using brute force and assuming that the agent has a fixed and discrete action set u at each state (e.g., turn rate command), the number of nodes at each level l of the decision tree is $|u|^l$ and each node should store the whole world knowledge because it is changing. This implies an exponential increase in time and space, making the infinite horizon optimization NP-hard [29].

[3] This type of optimization has been previously studied focusing in decentralization or in the receding horizon approach in papers such as [18,2,31,11], respectively.