Computers and Electronics in Agriculture xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

Task-driven active sensing framework applied to leaf probing

Sergi Foix*, Guillem Alenyà, Carme Torras

Institut de Robòtica i Informàtica Industrial, CSIC-UPC, Llorens i Artigas 4-6, 08028 Barcelona, Spain

A R T I C L E I N F O

Keywords:
Active perception
Next best view
Information gain
Search space reduction

A B S T R A C T

This article presents a new method for actively exploring a 3D workspace with the aim of localizing relevant regions for a given task. Our method encodes the exploration route in a multi-layer occupancy grid map. This map, together with a multiple-view estimator and a maximum-information-gain gathering approach, incrementally provide a better understanding of the scene until reaching the task termination criterion. This approach is designed to be applicable to any task entailing 3D object exploration where some previous knowledge of its approximate shape is available. Its suitability is demonstrated here for a leaf probing task using an eye-in-hand arm configuration in the context of a phenotyping application (leaf probing).

1. Introduction

The goal of task-driven exploration is to iteratively change the point of view so as to maximise the acquisition of information for solving a given task. We propose an algorithm that uses an information-gain criterion to compute the expected benefit of a set of candidate views, and combines it with other aspects, such as the proximity to the current view, to obtain the best next possible view at each iteration. Observe that this is a local approach, and that it cannot be globally optimal since each new position of the sensor only depends on the available information at each iteration. An optimal solution would require a complete and accurate model beforehand.

This article emphasises the following four main ideas:

1. A multi-layer occupancy map approach can naturally encode all the knowledge: each layer codifies relevant information that is semantically different (Section 3.3). Particularly for leaf probing tasks, the space occupied by the leaf and its surrounding clearance for allowing the tool to reach the leaf.
2. The importance of the termination criterion. Our representation, that explicitly represents the termination conditions in a specific occupancy layer, facilitates its definition and evaluation (Section 3.3).
3. Given a set of candidate viewpoints (Section 3.4), they can be evaluated using Information Gain ($\mathscr{IG}$), which can be easily defined and computed from the multi-layer representation (Section 3.5). A novelty in our proposal is that free space is also used for the $\mathscr{IG}$ computation.
4. The accurate characterization of the sensor used, in our case a time-of-flight camera (ToF) (Section 3.6), plays an important role in the definition and the computation of the $\mathscr{IG}$.

This article is an extended version of work published in Foix et al. (2015). We generalize our previous work by showing how the exploration models can be created depending on the characteristics of different tasks. Moreover, we demonstrate through validation experiments the robustness of our multi-layer $\mathscr{IG}$ criterion in conjunction with our frustum-based inverse sensor model, adding deeper explanations about the way to estimate the initial leaf pose.

2. Antecedents

Plant phenotyping studies the influence of environmental factors on the observable traits of plants. The success of such studies depends on the data extracted from a series of long-term monitoring experiments over a large number of plants under multiple environmental conditions. Measures can be obtained in two different scenarios. The first one includes regular fields and mobile sensors, either mounted on aerial vehicles using remote sensing techniques (Colomina and Molina, 2014), or on ground robots (Nielsen et al., 2012). Obviously, climate conditions cannot be controlled here. The second one uses greenhouses, where variables like temperature, humidity, and light, can be controlled. The common setup includes large greenhouses with several isolated zones, and conveyor belts that carry each plant from its sitting position to a measure chamber where a rich set of sensors takes measurements before returning them (Berger, 2013) (see Fig. 1). The throughput obtained in such installations is considerably high. However, sensors are located in a predefined position, mainly a top-view,

* Corresponding author.
E-mail addresses: sfoix@iri.upc.edu (S. Foix), galenya@iri.upc.edu (G. Alenyà), torras@iri.upc.edu (C. Torras).

https://doi.org/10.1016/j.compag.2018.01.020
Received 9 August 2017; Received in revised form 20 December 2017; Accepted 20 January 2018

(a) Typical greenhouse.[1] Image extracted from http://www.lemnatec.com/

(b) Monitoring chambers.[2] Image extracted from http://www.phenome-fppn.fr/

Fig. 1. Example of a modern plant phenotyping greenhouse. (a) Plants are kept into labelled pots over conveyor belts to easily monitor them when being transported from the greenhouse to the watering, nutrient delivery or monitoring chambers. (b) Plants get measured one by one in the different monitoring chambers.

that limits the possible photogrammetry tasks to be carried out, such as measuring the leaf length and rosette area (An et al., 2016), and extracting the plant 3D structure, either by using a depth camera (Chn et al., 2012), a stereo configuration (Aksoy et al., 2015), or a single RGB camera moving in a fixed direction (Jay et al., 2015). Note that these systems can not deal with hard occlusions. Another main limitation is the difficulty to measure or perform actions that require contact with the plant, such as chlorophyll measurement or the extraction of disk samples for DNA analysis (Alenyà et al., 2013).

Therefore, a major step forward is to provide the system with the ability to move its perceptual unit, so that it can naturally adapt to the characteristics of each plant (Bajcsy, 1988; Tsotsos et al., 1995). In the context of the European project GARNICS,[1] an active perception system was proposed to overcome this weakness that involves ToF camera and a probing tool mounted on the end-effector of a robot manipulator (Alenyà et al., 2014).

Classically, the *task* (usually implicit) in sensor path planning has been precise 3D object's surface reconstruction (Scott et al., 2003; Kriegel et al., 2015) and also object recognition (Roy et al., 2004). Less commonly, sensor path planning has also been used to optimally segment particular object characteristics (Madsen and Christensen, 1997; Saxena et al., 2008) and to exploit sensor features for simplifying occlusions detection, formerly using a laser (Maver and Bajcsy, 1993) and more recently using a ToF sensor (Foix et al., 2011).

Information gain has widely been used as *viewpoint selection criterion* in classical 3D modelling (Curless and Levoy, 1996; López-Damian et al., 2009). Newer approaches compute the expected information gain using a discrete approximation of the sensor's field of view based on pixel's ray-tracing techniques. They can be single-resolution (Potthast and Sukhatme, 2014) or multi-resolution (Vasquez-Gomez et al., 2013). In contrast, our proposal uses the complete pixel's frustum for a more accurate computation. Although a little bit slower, our approach assures to not miss any space between rays independently of the resolution of the octree. This may be negligible in large indoor/outdoor mapping (Bissmarck et al., 2015), but it is crucial in our short range application.

An important peculiarity of our approach is that, since ToF sensors uncertainty is not uniform, a precise calibration of the sensor is required to adequately model the image acquisition process and thus improve the $\mathscr{IG}$ computation. In previous approaches, the sensor's uncertainty is considered uniform for all the acquired points. Up to our knowledge, the different uncertainty present in each pixel has not been used before for the precise computation of the $\mathscr{IG}$. Our model assures a good approximation of the sensor's uncertainty distribution based on both pixel's location and measured distance.

The most common way for representing the 3D space is by using occupancy grid maps (Bissmarck et al., 2015; Vasquez-Gomez et al., 2015). An interesting idea for 2D robot navigation was proposed by Lu et al. (2014), where a multilayer costmap approach was used to separate different semantic information on different layers. In this article, we extend this idea to 3D and proposed a new multi-layered occupancy map to separate the state map, the obstacle avoidance map and the termination criterion.

Finally, it is crucial to define a *termination criterion* that is relevant to the task. The most used ones have been the ratio of visible and occluded information in the scene (Banta et al., 2000), and different measures of the completeness of the model (Kriegel et al., 2015; Torabi and Gupta, 2012). However, completeness is usually non-intuitive to define. Vasquez-Gomez et al. (2015) proposed also to take into account if a mobile robot could not find a path for any of the remaining candidate views. In contrast, in this work we link the task's goal and the termination criterion by explicitly defining specific regions of interest (ROIs), one or more, in a unique 3D layer map.

3. Leaf probing

The main idea behind our method is that, based on the previous knowledge of an exploration task, we can pre-establish a 3D occupancy map and a set of candidate views that, all together, indirectly serve as a guide to a Next-Best-View (NBV) planner for solving the task (Fig. 3). On the one hand, the 3D occupancy map is used for locating a set of possible regions of interest for the given task. On the other hand, the set of candidate views is used for reducing the dimensionality of the gaze space while ensuring a complete coverage of the regions of interest.

The approach includes three main steps (see Fig. 4):

1. Selection of the target leaf from a plant
2. Exploration of this leaf to gather enough relevant information for the task at hand
3. Effective execution of the task

The first step is usually specified by a botanical expert that defines a criterion to choose the leaf, for example the biggest one, or the *i*-th leaf from the stem. The last step, the measuring action, is been carried out as proposed in Alenyà et al. (2013). This article focuses on a method for solving the second step.

In the following sections we will explain every module of our approach in detail. Due to its high complexity, the task of leaf probing is the one used for illustration purposes and real experimental evaluation. Observe how this task does not require to have a complete leaf's model to accomplish its goal. Instead, only specific regions in the leaf's

[1] http://www.garnics.eu/.