Available online at www.sciencedirect.com
ScienceDirect
IFAC-PapersOnLine 49-16 (2016) 085–089

Semantic Mapping of Orchards ⋆

Cheng Peng * **Pravakar Roy** * **James Luby** ** **Volkan Isler** *

* *Computer Science & Engineering, University of Minnesota, Minneapolis, MN 55455 {peng0175,proy,isler}@cs.umn.edu*
** *Horticultural Science, University of Minnesota, St Paul, MN 55108 lubyx001@umn.edu*

Abstract: We present a method to construct a semantic map of an apple orchard using a LIDAR and a camera rigidly attached to each other. The system is able to capture the map as a standalone sensor which is light-weight and can be mounted on a variety of platforms.
At the geometry level, we present a new method to associate image features captured by the camera with 3D points captured by the LIDAR. We then use this method to register 3D point-clouds onto a common frame. We show that our association method yields superior registration performance compared to common methods which work in indoor or urban settings.
At the semantic level, the apples are identified as distinct objects. Their locations and diameters are extracted as relevant attributes. As an example, a semantic map of an orchard row is constructed.

© 2016, IFAC (International Federation of Automatic Control) Hosting by Elsevier Ltd. All rights reserved.

Keywords: computer vision; agriculture; image reconstruction; image recognition; object recognition

1. INTRODUCTION

Robotics and sensing technologies are finding increasing application in agricultural sciences and practices. In these applications, data collected by sensors is turned into information at various levels. At a very basic level, the sensor data acquired from a mobile platform is mapped directly to relevant information. A typical example is the mapping of aerial images to Normalized Difference Vegetation Index (NDVI) tables based on multi-spectral color and intensity data. At the next level, geometric maps are built as an intermediate step. From these maps, information such as crop height can be directly estimated, for example for phenotyping applications (See Li et al. (2014) for an overview).

In order to turn sensor data into actionable information, what is ultimately needed is a "semantic map" in which relevant objects (plants, fruits, branches, etc.) and their relevant attributes (size in three dimensions, color across multiple spectra, etc.) have been identified. In this paper, we present a system which can acquire both a geometric map and a semantic map of an orchard using only camera and LIDAR images (without additional information from GPS or inertial sensors).

Our main technical contribution for the geometric level is a novel method for accurately associating LIDAR points with image features which allows for the quick registration of LIDAR point clouds that can be used to reconstruct the plant stature, shape, and canopy density. Rapid, efficient three-dimensional plant canopy reconstruction has applications for phenotyping in research and informing plant canopy management practices such as pruning and fruit thinning. Geometric reconstruction is a challenging problem because a close-up view of a tree has numerous discontinuities. Therefore, two pixels close to each other in the image can be far in space. At the semantic level, our system is capable of identifying apples and extracting information about their locations, diameters and count which can provide rapid phenotpying for fruit yield, yield efficiency, and fruit distribution in the canopy relative to canopy density. We show an example of such semantic map constructed from imagery collected in an apple orchard. In the next section, we start with an overview of the related work.

2. RELATED WORK

Semantic mapping can be very useful in agricultural applications by providing relevant information to farmers, growers or biologists. For speciality farms in particular, it is important to locate individual trees and fruits, and estimate their sizes so as to calculate useful yield parameters. The topic of semantic mapping has been studied extensively in the context of mobile robots (Kostavelis and Gasteratos (2015)). In this paper, we restrict ourselves to work directly relevant to agriculture. Weiss and Biber (2010) presented a semantic place classification system for outdoor agricultural robots. In contrast to our approach they used high resolution RTK-GPS devices for constructing the maps. Zhang et al. (2014) presented a landmark-based method for creating a local map of the environment using a LIDAR. Their system is mounted on a ground vehicle. They use the vehicle position determined by encoder odometry to register the LIDAR points. In contrast, we utilize the image features from our LIDAR - camera system to register the LIDAR points. Das et al. (2015) presented a sensor suite consisting of a laser range scanner, multi-spectral cameras, a thermal imaging camera, and

⋆ This work is supported in part by NRI Award 1525045, RI Large Award 1111638, NSF Award 1317788, USDA Award MIN-98-G02 and the MnDrive initiative.

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.2016.10.016

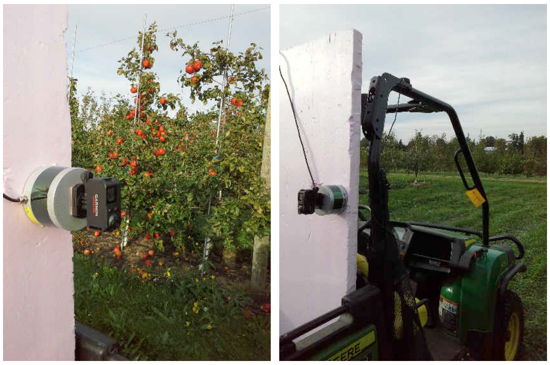

Fig. 1. **Left:** Camera+LIDAR system used for data capture **Right:** The system mounted on a ground vehicle.

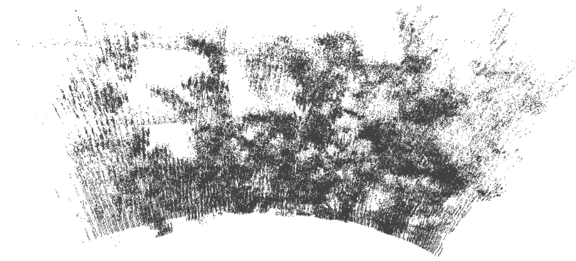

Fig. 2. Iterative Closest Point registers frames of LIDAR points in a curved way.

navigational sensors. They presented techniques to extract four key data products - plant morphology, canopy volume, leaf area index, and fruit counts using the sensor suite. They build a map of the rows using the thermal sensor and registered the point clouds using the navigational sensors. Bargoti et al. (2015) presented a pipeline for trunk detection in trellis structured apples orchard. Similar to the previous approaches, they also depend on odometry and GPS for registering the LIDAR data. In contrast, we register the LIDAR data completely based on visual features and focus on extracting semantic information such as fruit location and size. In the next section, we start with a description of our system.

3. SYSTEM AND DATA ACQUISITION

A camera (GARMIN Virb XE) and LIDAR (Velodyne VLP-16) fusion system is used for data collection (Fig 1). The camera is calibrated using the method of Zhang (2000). As the LIDAR reflection is not visible on the camera image, the relative pose of the two sensors needs to be calculated. The method of Park et al. (2014) is used to calibrate the camera and LIDAR. The images and point clouds are synchronized using the time stamps provided by both devices.

The system is powered with a dedicated power pack (Qi-infinity Powergrid 34,200mAh). The system is light enough (2.16 lbs without battery and 5.37 lbs with battery) so that it can be mounted on various platforms. It is also stand-alone and does not need to be interfaced with additional sensors on the platform. This makes data acquisition very convenient.

In our case, the sensor system was mounted on a ground vehicle that drives in between orchard rows while the system captures images and the LIDAR points of the trees. The ground vehicle is driving at a relatively constant speed (5mph) to avoid jerky motion that deteriorates the image and LIDAR data quality. The slow motion also ensures enough feature correspondences among images to recover the motion robustly.

4. GEOMETIC RECONSTRUCTION

Since our system acquires 3D point clouds with LIDAR, data acquisition at the geometry level is performed directly.

However, to register any two point clouds captured by the LIDAR onto a common frame, sensor motion needs to be estimated. This *registration* problem can be formulated as follows:

Suppose the sensor system moved from point A to B, where the respective LIDAR centers are at l_1 and l_2. The resulting image ${}^{l_1}\mathbf{I}, {}^{l_2}\mathbf{I}$ and laser points ${}^{l_1}\mathbf{P}, {}^{l_2}\mathbf{P} \in \mathbb{R}^3$ are with respect to the local coordinate frames l_1 and l_2, which satisfies the following equation:

$$ {}^{l_1}\mathbf{P} = {}^{l_1}_{l_2}\mathbf{R}\,{}^{l_2}\mathbf{P} + {}^{l_1}\mathbf{t}_{l_2} + \mathbf{N} \quad (1) $$

where ${}^{l_1}_{l_2}\mathbf{R} \in \mathbb{R}^{3\times3}$, ${}^{l_1}\mathbf{t}_{l_2} \in \mathbb{R}^3$ are the rotation matrix and translation vector respectively and $\mathbf{N} \in \mathbb{R}^3$ is the noise vector. The goal of registration is to compute the transformation matrix:

$$ {}^{l_1}_{l_2}\mathbf{T} = \begin{bmatrix} {}^{l_1}_{l_2}\mathbf{R} & {}^{l_1}\mathbf{t}_{l_2} \\ 0 & 1 \end{bmatrix} $$

Without noise, if we can find four pairs of corresponding LIDAR points between two frames, we can solve for the transformation matrix directly (assuming that any three points are not collinear). However, in practice, since the sensors are moving forward, the LIDAR data is sparse and the scene has many occlusions, it is nearly impossible to acquire the same point in two consecutive images.

When the correspondences between point clouds are unknown, the most common technique for aligning two sets of point clouds is the *Iterative Closest Point*(ICP) method (Besl and McKay (1992)). ICP iteratively estimates point correspondences and transformation matrix until a convergence criterion is met. For our data set, ICP failed as shown in Fig 2. In each frame, LIDAR points are relatively sparse and recorded vertically. The vehicle is moving parallel to the scene which results in fan-shaped data-points. Maximizing alignment with ICP results in curved registration.

We observed that we can find accurate image correspondences using robust features estimator such as SIFT (Lowe (2004)). Therefore, an alternative approach is to use image features to find correspondences across frames. If we can associate depth values to these features then we can compute the transformation $\mathbf{T}$ directly. In the next section, we turn our attention to this problem.

Download English Version:

https://daneshyari.com/en/article/5002396

Download Persian Version:

https://daneshyari.com/article/5002396

Daneshyari.com