Computers and Electronics in Agriculture 128 (2016) 172–180

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

A novel tree trunk detection method for oil-palm plantation navigation

Mohammed Ayoub Juman *, Yee Wan Wong, Rajprasad Kumar Rajkumar, Lay Jian Goh

Department of Electrical and Electronic Engineering, The University of Nottingham, Malaysia Campus, Jalan Broga, 43500 Semenyih, Malaysia

A R T I C L E I N F O

Article history:
Received 30 May 2016
Received in revised form 1 August 2016
Accepted 2 September 2016
Available online 14 September 2016

Keywords:
Tree trunk detection
Computer vision
Oil palm plantation
Microsoft XBOX KINECT
Mobile robot

A B S T R A C T

This paper presents a novel tree trunk detection algorithm that uses the Viola and Jones detector along with a proposed pre-processing method, combined with tree trunk detection via depth information. The proposed method tackles the issue of the high false positive rate when the Viola and Jones detector is used on its own, due to the low contrast between tree trunks and the surrounding environment. The pre-processing method uses colour space combination and segmentation to eliminate the ground not covered by trees from the images and feeding the resulting image into a cascade detector for identifying the location of the trunks in the image. Depth information is obtained via the use of the Microsoft KINECT sensor to further increase the accuracy of the detector. Our proposed method had better performance when compared to both Neural Network based and Support Vector Machine based detectors with a detection rate of 91.7% and had the lowest false acceptance rate out of other detectors, including the original Viola and Jones detector. The performance of the proposed method was also tested on live video feeds with the use of a robot prototype in an oil-palm plantation, which proved the high accuracy of the method, with a 97.8% detection rate. The inclusion of depth information resulted in more accurate detections during low levels of light and at night, where reliance on pure depth information resulted in a 100% detection rate of tree trunks within the range of the sensor.

1. Introduction

Automation in agricultural processes is becoming more of a necessity than a luxury in recent times. With the increase in demand for manpower and the decline in seasonal labour, producers have no choice but to modernise their practices to avoid reduction of yields as well as overall profits. Automated agricultural vehicles that cater to varied tasks such as weeding, pesticide application, fruit harvesting and even tree pruning are on the rise to meet the needs of the ever growing development in the field.

With the development of sensors and machine vision techniques, several agricultural vehicles have been upgraded for automated navigation, harvesting and mapping of various sites. Obtaining readings of the working region is a key criterion for autonomous functionality, as this opens the way for object detection and mapping, both crucial processes to provide references for navigation. Cameras provide a low cost solution for obtaining information of the immediate environment by means of intensity and colour data, while range sensors such as sonar and lasers are used to obtain both 2 and 3 dimensional depth data.

The semi structured nature of orchards and plantations make autonomous navigation possible with the use of basic machine vision techniques and sensors. Using a planar laser scanner on a mobile robot, row by row navigation can be carried out. In these cases, the scanner is able to obtain range information of the trees on both sides of the row, enabling the mobile robot to centre itself and navigate to the end of the given row. Bayar et al. (2015) used a planar range finder to align their robot to orchard rows, focusing more on the row like arrangement than individual trees. Gimenez et al. (2015) used similar methods, although their system required prior knowledge of the absolute points of the end trees of the rows.

Although obtaining range information from planar sensors is useful for knowing the distances to obstacles and trees, they do not provide sufficient data for creating a good reconstruction of a given area by themselves. In certain cases, planar data is not sufficient to detect obstacles, leading to the need of more range data to make accurate assessments. As a result, the use of 2D LIDAR sensors to obtain 3D point clouds has increased in popularity in recent years. Given the high number of data points for each square meter, complete scene reconstruction is entirely possible, leading to higher chances of accurate detections. Sanz et al. (2013) used a 2D terrestrial LIDAR scanner for characterising tree crops by scanning each row on both sides via a tractor mounted sensor to obtain

* Corresponding author.

E-mail addresses: kecx4mam@nottingham.edu.my (M.A. Juman), YeeWan.Wong@nottingham.edu.my (Y.W. Wong), Rajprasad.Rajkumar@nottingham.edu.my (R.K. Rajkumar), kecy2gla@nottingham.edu.my (L.J. Goh).

http://dx.doi.org/10.1016/j.compag.2016.09.002
0168-1699/

two 3D point clouds. The obtained data enabled them to calculate accurate leaf area density of the present trees. Similar work was done by Rosell et al. (2009) to obtain large amounts of plant information, giving a high correlation coefficient of 0.976 between trees reconstructed using data points and manual measurements. Forsman and Halme (2005) displayed the use of dense range images to reconstruct trees by representing them as cylindrical columns in scans of large scale environments. Their reasoning being that accurate mapping would enable virtual testing of forest working machines. Zhang et al. (2014) used a 2D LIDAR to generate maps of orchards to simplify autonomous navigation. Their system required marked posts as artificial landmarks for better judgement of row ends, producing results with only 19–26 cm mean distance error.

Other research has used machine vision techniques for identification in agricultural environments, even without the use of sensors. Auat Cheein et al. (2011) tackled precision agriculture mapping with the use of support vector machines (SVMs) for detecting olive stems from images obtained via a monocular vision system. Once the trees were detected by the system, a laser range sensor was used to give the distance of the stem to the robot. The absence of weeds and a clear contrast between the ground and the trunks made the use of a pure vision based system suitable for this scenario. Lu and Rasmussen (2011) used an omnidirectional camera and contrast templates to detect tree trunks by relying on the attribute where trunks appear as straight lines with a high contrast at both sides. Shao et al. (2014a) used colour features of the trunk from the $L*a*b$ colour space as well as shape features by means of the Hough transform to detect trees. Shao et al. (2014b) furthered their work with the addition of a back propagation neural network to segment the colour marks from the $L*a*b$ colour space before performing the Hough Transform. Torres-Sospedra and Nebot (2014) used artificial neural networks for the detection of weeds in orange groves enabling their system to detect weeds with a very high performance rating.

Though there are several systems that make use of sensors or cameras by themselves, certain environments/situations require more data than either can provide individually. The solution for these cases is to make a system that combines the information obtained from both sensors and cameras so that the lack of information from one input is made up by the other. Jaakkola et al. (2010) worked with a multi sensorial system for mobile mapping, though their method was unable to distinguish between trees and other pole type objects. Zhong et al. (2013) used sensors along with a CCD panoramic camera to incorporate texture information into the reconstruction of the 3D model of the environment. Shalal et al. (2015a) were able to detect trunks with 96.6% accuracy with the use of a camera and laser scanner, similar to the previous systems, highlighting the advantage of using multiple input devices for increased detection accuracy. They predicted that although their system had high accuracy in the orchard, it might have poor results if the trunks were occluded with leaves/fauna. They furthered their work by using the tree detector for mapping and localisation by using the trunks as key points in the map (Shalal et al., 2015b).

With the introduction of the Microsoft XBOX KINECT sensor that incorporates both vision and depth inputs, the possibilities of its usage for research purposes have increased over the recent years. Though mainly used for indoor applications due to the interference of natural light with the infrared light used by the depth sensor, there are a few cases where it has been used in outdoor/agricultural settings. Nissimov et al. (2015) used the KINECT in a greenhouse environment for obstacle detection via the use of depth and colour information. The slope of the environment ahead was calculated using depth data and if the slope was too large to be classified as ground, colour and texture features were used to classify the obstacle into various categories. Though not in an agricultural setting, Arnay et al. (2016) used the KINECT for obstacle detection outdoors for autonomous vehicle navigation. They managed to circumvent the problem of natural light by placing the KINECT close to the ground and tilting it to face at a downward angle, reducing the chances of interference and focusing on the obstacles at close quarters.

This paper deals with the detection of tree trunks in an oil-palm plantation to enable further work to be done in creating a navigation system capable of operating within the area autonomously. Though several works have been done in orchards and other plantation settings, this is the first to focus on oil-palm plantations with robot navigation in mind. In the plantation, though the environment is relatively structured, inconsistent gaps between trees and landscape variations lead to differing distances between plots and the number of trees within them. These factors, along with the size of the land to be covered as well as the changing landscape due to bush growth and fallen trees/branches, eliminate the use of preprogrammed methods. This leads to the need of an automated system capable of detecting trees as key points for navigation as they are the main point of interest. The focus of this work was to produce a detector with high accuracy while keeping the costs as low as possible to make it a more realistic prototype that could be used by regular plantation owners.

We propose a novel pre-processing method focused on reducing the image size to the area we assumed to have trees which is then fed into the detector for further processing. The method works by using colour space combination to enhance the contrast between the ground and other objects and then cropping the image to remove ground areas. Varying lighting conditions and the low contrast between tree trunks and the background in visual information made the use of other sensory information a necessity for improving the accuracy of the system. In keeping to low costs, the KINECT sensor was used for obtaining video input as well as depth information, instead of having two separate systems as inputs. The detector used on the video input was created by Viola and Jones (2001), chosen for its cascade nature, which eliminates unlikely inputs in the early stages of the cascade. Furthermore, the depth data from the KINECT was used in tandem with the detector to enable more accurate detections. This paper is the first to investigate the use of KINECT's depth data in tree trunk detection performance in an oil-palm plantation environment. The proposed system was tested out in the field with the use of a mobile robot designed to navigate and collect data in an oil-palm plantation.

The remainder of the paper is divided as follows. Section 2 describes the plantation that the work was focused on and the properties as well as challenges expected from such an environment. The methodology with which the robot was designed is covered in Section 3 while plantation data obtained and the workings of the proposed system are mentioned in Section 4. Results are displayed and discussed in Section 5. The work done is concluded in Section 6.

2. Study of the plantation

The Balau Estate oil-palm plantation in Broga, Malaysia, was used to collect data as well as test the detector. In the plantation, trees were planted in square plots of 20×20 trees, with a distance of 8.8 m between each tree, where every 3 adjoining trees form a triangle. The layout was similar to arranging them in parallel rows, with a distance of about 7.8 m between each row (Fig. 1(a)). Branches that are cut as well as fallen/felled tree trunks are placed in the middle of rows so that they may decompose and add nutrients to the soil. Field paths exist between each square plot, where