ISPRS Journal of Photogrammetry and Remote Sensing 88 (2014) 128–146

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Indoor and outdoor depth imaging of leaves with time-of-flight and stereo vision sensors: Analysis and comparison

Wajahat Kazmi [a,*], Sergi Foix [b,1], Guillem Alenyà [b,1], Hans Jørgen Andersen [a,2]

[a] Department of Architecture, Design and Media Technology, Aalborg University, Sofiendalsvej 11, 9200 Aalborg, Denmark
[b] Institut de Robòtica i Informàtica Industrial, CSIC-UPC, Llorens Artigas 4-6, 08028 Barcelona, Spain

ARTICLE INFO

Article history:
Received 30 May 2013
Received in revised form 1 October 2013
Accepted 23 November 2013
Available online 31 December 2013

Keywords:
Leaf imaging
Depth
Exposure
Time-of-Flight
Stereo vision
Sunlight

ABSTRACT

In this article we analyze the response of Time-of-Flight (ToF) cameras (active sensors) for close range imaging under three different illumination conditions and compare the results with stereo vision (passive) sensors. ToF cameras are sensitive to ambient light and have low resolution but deliver high frame rate accurate depth data under suitable conditions. We introduce metrics for performance evaluation over a small region of interest. Based on these metrics, we analyze and compare depth imaging of leaf under indoor (room) and outdoor (shadow and sunlight) conditions by varying exposure times of the sensors. Performance of three different ToF cameras (PMD CamBoard, PMD CamCube and SwissRanger SR4000) is compared against selected stereo-correspondence algorithms (local correlation and graph cuts). PMD CamCube has better cancelation of sunlight, followed by CamBoard, while SwissRanger SR4000 performs poorly under sunlight. Stereo vision is comparatively more robust to ambient illumination and provides high resolution depth data but is constrained by texture of the object along with computational efficiency. Graph cut based stereo correspondence algorithm can better retrieve the shape of the leaves but is computationally much more expensive as compared to local correlation. Finally, we propose a method to increase the dynamic range of ToF cameras for a scene involving both shadow and sunlight exposures at the same time by taking advantage of camera flags (PMD) or confidence matrix (SwissRanger).

1. Introduction

In agricultural automation, 2D imaging has addressed a variety of problems, ranging from weed control (Slaughter and Giles, 2008) and disease detection (Garcia et al., 2013) to yield estimation (Nuske et al., 2011), inter-plant space sensing (Tang and Tian, 2008) and structural analysis (McCarthy, 2009), to name a few. But most of these tasks are either large scale analysis or they tend to deal with simpler plant canopies, for example, at early growth stages (Astrand and Baerveldt, 2004). The reason is obvious; when looking inside plant canopy, 2D imaging is not robust to occlusion of plant organs such as overlapping leaves and branches.

To address this problem, 3D imaging is a common solution. Among the most noticeable applications of 3D vision are the construction of dense models for simulation of plant structures (Takizawa et al., 2005) and for estimating 3D properties of plant canopies (Chapron et al., 1993; Preuksakarn and Boudon, 2010; Santos and Oliveira, 2012). If not obvious, then at least an ambiguous difference between the application domains of 2D and 3D imaging in agriculture can be observed. 2D has been successfully applied for outdoor and 3D for indoor applications and large scale analysis in outdoor scenario such as navigation in the field (Kise and Zhang, 2008).

The reason for this gap is threefold; firstly, plants have complicated free form, non-rigid structures that cannot be approximated by simple geometrical shapes making it necessary to observe minute details and hence placing stringent demands on the quality and the efficiency of 3D imaging technology. Secondly, the huge variations in outdoor illumination (sunny, partially cloudy, overcast, shadow), which can change the perceived shape of objects to a large extent and that even constrains 2D imaging. Thirdly, the technology for 3D data acquisition is largely designed for indoor applications and exporting it to outdoor scenario either limits the scope or makes the system too complex to be practical. For example, (Biskup et al., 2007) used stereo vision for only measuring leaf inclination for outer leaves of plant canopies under outdoor lighting and (Nakarmi and Tang, 2010) used a ToF camera for plant space measurement covering the view from sunlight otherwise the sensor saturates. In general, any such approach for 3D analysis

* Corresponding author. Tel.: +45 9940 7156.
E-mail addresses: wajahat@create.aau.dk (W. Kazmi), sfoix@iri.upc.edu (S. Foix), galenya@iri.upc.edu (G. Alenyà), hja@create.aau.dk (H.J. Andersen).
[1] Tel.: +34 9340 54261.
[2] Tel.: +45 9940 8834

0924-2716/$ - see front matter © 2013 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS) Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.isprsjprs.2013.11.012

is focused at a particular application and cannot be easily adapted for obtaining slightly different measurements.

Even after all the shortcomings, 3D sensing is vital. Plant phenotyping facilities require accurate depth measurements of plant organs (such as, leaf count/angles/areas, plant height or sampling points on specific sections of a plant) either for classification of large varieties of plants produced due to experimental variations of environmental factors (Van der Heijden et al., 2012) or robotic manipulation such as for measuring chlorophyll content in leaves (Alenya et al., 2011) or automated fruit picking (Jimenez et al., 2000). In field operations, it has great potential in precision agriculture for reducing the amount of herbicides as 3D data can help in not just improved recognition and localization of weeds by resolving occlusion but also in estimation of volume of the infestation, thereby enabling deployment of optimal amounts of chemicals (Nielsen et al., 2004;(Kazmi et al., 2011).

Recently, Fiorani et al. (2012) discussed state-of-the-art technologies in use for biological imaging and pointed out that in depth knowledge is required regarding physics of the sensors and parameters of software/algorithms used in order to benefit optimally. This is a bottleneck in agricultural automation because the objects (plants) pose one of the most demanding tests to image acquisition and machine vision. Systems optimized for man made structured environments are not optimal for the natural setup of agriculture. Limitations of imaging system combined with environmental factors make agricultural imaging a complex puzzle to solve. Therefore, it is important to segregate environmental factors and evaluate the sensor performance w.r.t to each one, individually.

One of the most important factors is light, both indoor and outdoor. Lighting must be diffused to reduce errors. Under outdoor conditions, various shading arrangements have been used to cater for this or else experiments are performed on days with overcast (Frasson and Krajewski, 2010). But the problem arises when introducing a shade makes the system either too complicated, such as, in weed detection (Piron et al., 2011) or sunlight is unavoidable, for example, to understand the effect of lighting variations on the plant canopies (Van der Zande et al., 2010), to track the diurnal/nocturnal movement of the leaves (Biskup et al., 2007) or with changing positions of the sun (Van Henten et al., 2011). In such cases, the exposure of the 3D imaging system must be either robust to variation in ambient illumination or at least tangible, somehow. The effect of ambient illumination on the camera response varies with the type of sensor used.

1.1. Common 3D data acquisition techniques and challenges

The most widespread method of acquiring 3D data is stereo vision. But it has a big set of problems. Stereo correspondence and depth accuracy vary with the type of algorithm used. Local correspondence algorithms are efficient but less accurate than global ones which could be, computationally, very expensive. Besides, performance is adversely affected by lack of surface texture of the object and specular highlights.

Among the active sensing technologies, structured light projection and laser range scanners are used for creating accurate and detailed 3D models, but such systems are usually expensive and complex. On the other hand, structured light has interference issues under sunlight and laser scanners include mobile parts that cause longer imaging times. Although, new low cost versions of structured light cameras have recently appeared, they still offer low resolution and are highly sensitive to outdoor lighting (such as RGBD cameras e.g. Microsoft Kinect[3]).

On the other hand, recent advances in the ToF based range sensors have revolutionized the industry and several brands of off-the-shelf 3D cameras are available in the market. They use near infrared (NIR) emitters and generally produce low resolution depth images. However, a gradual increase in sensor resolution has been observed over the last few years. ToF cameras produce high frame rate (up to 50 fps) depth images and therefore, are highly suitable for real-time applications. But the problem of lack of performance under sunlight, still remains i.e. these sensors are guaranteed to work only in indoor environments. Some of the ToF cameras have an on-board background illumination rejection circuitry such as PMD (Möller et al., 2005), but with varying performance under sunlight depending on the operating range and the power of NIR emitters.

The challenge in ToF cameras is to find a suitable integration time (IT: a controllable parameter related to the time the sensor spends integrating the returned signal) according to the ambient illumination because a different calibration has to be applied for each IT and the calibration is a costly process. For stereo vision, the challenge is the performance and accuracy of the correspondence algorithm and the effects of ambient illumination on the accuracy of disparity map. Fig. 1 shows a comparison of working principle and Fig. 2 of the data processing pipelines for both stereo vision and ToF technologies.

In our previous work, we have evaluated the performance of one ToF camera for close range leaf imaging (Kazmi et al., 2012). But every ToF camera has different sensor properties and robustness against background illumination. A qualitative comparison of the response of several different ToF cameras with stereo vision under indoor/outdoor illumination conditions, particularly for agricultural purposes, is not available in literature. Such sensor characteristics would be very helpful for analyzing the performance of these sensors and weighing the cost of making a choice.

1.2. Objective

In this article, our objective is to estimate and compare the response of ToF and stereo vision sensors for depth imaging of leaves using some of the commonly used cameras. We will first review their current applications in agriculture. Since a lot of literature has addressed resolution and accuracy of stereo vision (e.g. Scharstein and Szeliski, 2002; Kytö et al., 2011) we will only provide a short insight into the precision of ToF cameras.

We will introduce some metrics for qualitative evaluation of the depth data. We also propose a method for obtaining the most suitable camera configurations for imaging under different illumination conditions. The method is based on observing the trends in camera precision and detecting the non-linearities in the amplitude. Additionally, we show that for ToF cameras, using this

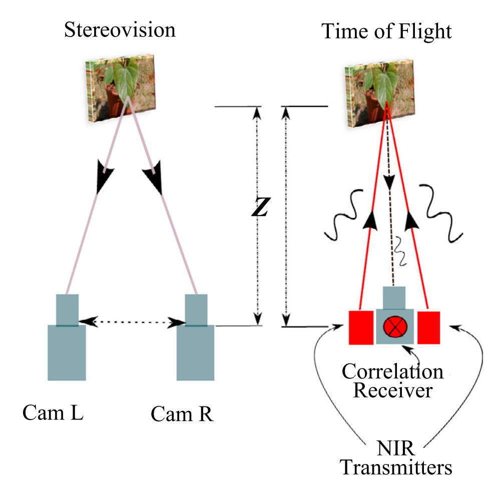

Fig. 1. Comparison of stereo and ToF techniques.

[3] http://www.microsoft.com/en-us/kinectforwindows.