Scientia Horticulturae 228 (2018) 187–195

Contents lists available at ScienceDirect

Scientia Horticulturae

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

Research Paper

A comparative evaluation of combined feature detectors and descriptors in different color spaces for stereo image matching of tree

Ayoub Jafari Malekabadi, Mehdi Khojastehpour*, Bagher Emadi

Department of Biosystems Engineering, Ferdowsi University of Mashhad, Mashhad, Iran

ARTICLE INFO

Keywords:
Tree
Depth map
Image matching
Feature detector
Feature descriptors
Color spaces

ABSTRACT

Tree canopy geometric characteristics are directly related to tree growth and productivity, and this information has been used to predict yield, fertilizer application in citrus crops, water consumption or biomass. So a 3D model and a depth map of the tree can be useful. One method of creating 3D model is stereo vision technique. The comparison and performance assessment of different combinations of feature detectors and descriptors is very important in this technique. In this study, the performance of 12 combinations of well-known detectors and descriptors including BRISK with SURF, BRISK with BRISK, BRISK with FREAK, Harris with SURF, Harris with BRISK, Harris with FREAK, SURF with SURF, SURF with BRISK, SURF with FREAK, MSER with SURF, MSER with BRISK, and MSER with FREAK was evaluated. The included color spaces were HSV, H, YCbCr, Y, NTSC and RGB. The performance comparison for each combination was carried out in terms of precision and recall values using stereo image pairs of tree. It was observed that the largest number of keypoints were detected by MSER and SURF detectors almost in all possible spaces. Best combinations used from SURF detector and descriptor when the precision and recall results were considered. RGB and Y spaces were the best spaces to implement combinations. The combination of SURF with SURF, SURF with FREAK, and HARRIS-SURF were found to be preferable.

1. Introduction

The structural aspects of a canopy are crucial at different levels (individual tree, crops, forest and ecosystems) (Phattaralerphong and Sinoquet, 2005). Most of the research conducted to date are related to forest areas (Lefsky et al., 2002; Parker et al., 2004; Maas et al., 2008; Kushida et al., 2009). However, in the field of agriculture, obtaining three-dimensional (3D) models of trees and plantations introduces an immense and novel field of applications. The geometric characterization of trees are both relevant and complex tasks (Sanz-Cortiella et al., 2011a,b). Canopy characteristics supply valuable information for tree management reducing production costs and public concerns about environmental pollution. Thus, there is a whole range of key agricultural activities including pesticide treatments, irrigation, fertilization and crop training which depend largely on the structural and geometric properties of the visible part of trees (Rosell and Sanz, 2012).

At present, the research groups are conducting research to a variety of non-destructive techniques for the measurement of the tree canopy structural characteristics such as volume, foliage and leaf area index. This can be achieved by different detection approaches. The use of ultrasonic sensors (Giles et al., 1988; Zaman and Salyani, 2004; Zaman and Schumann, 2005; Solanelles et al., 2006), as well as digital photographs (Phattaralerphong and Sinoquet, 2005; Leblanc et al., 2005), laser sensors (Naesset, 1997a,b; Aschoff et al., 2004; Van der Zande et al., 2006; Rosell et al., 2009a,b), stereo images (Andersen et al., 2005; Rovira-Mas et al., 2005; Kise and Zhang, 2006), light sensors (Giuliani et al., 2000), high-resolution radar images (Bongers, 2001) or high-resolution X-ray computed tomography (Stuppy et al., 2003) offer innovative solutions to the problem of structural assessment in trees (Rosell and Sanz, 2012; Jafari Malekabadi et al., 2016). For example, Giles et al. (1987, 1988, 1989a,b) discussed the use of the ultrasonic sensors to measure canopy volume in peach and apple trees. They developed using this technique to improve the process of pesticide application. The measurement system was based on a three ultrasonic sensors mounted at different heights of an air-blast orchard sprayer. The results showed pesticide saving up to 52% in apples.

Despite noted researches, little research has been done on tree modeling by machine vision. In particular, vision-based measurement methods are nondestructive and an effective way to determine external plant features (Yeh et al., 2014). Stereo vision is a method for the extraction of 3D information from digital images in this field. The most important problem is the corresponding points and matching them in

* Corresponding author.
E-mail address: mkhpour@um.ac.ir (M. Khojastehpour).

http://dx.doi.org/10.1016/j.scienta.2017.10.030
Received 18 June 2017; Received in revised form 13 October 2017; Accepted 16 October 2017
Available online 19 October 2017
0304-4238/

Table 1
Properties of images.

Property	Digital zoom	Flash mode	35 mm focal length	Focal length (mm)	F-stop	Resolution (dpi)	Dimensions (Pixels)
Value	0	No flash	24	4	f/7.4	300	2447 × 3264

this method. The well-known detectors and descriptors that receiving most citations are ORB, SIFT, SURF, BRISK, FREAK, HARRIS, FAST and MSER (Llorens et al., 2011). It is impossible to obtain best accuracy as well as best security in a minimum computation-time at the same time. Therefore, we have to make concessions for the sake of selection an optimal feature detection method with respect to the task performed.

Some feature detection methods are compared in several studies. Six feature descriptors have chosen to make a comparison: SURF, ORB, BRIEF, BRISK, SIFT and SU-BRISK by Peng (2012). Also, a comparative analysis of three binary descriptors (ORB, BRIEF and BRISK) by concentrating on well-known detectors (ORB, MSER, SIFT, SURF, FAST and BRISK) is carried out in terms of effects of various geometric and photometric transformations (Heinly et al., 2012). In a study on comparison of low level feature extraction algorithms (El-gayar and Soliman. 2013), the performance of FAST-SIFT (F-SIFT) feature detection methods have compared in case of blur, illumination and scale changes, rotation and affine transformations. In another study the comparison analysis between SIFT and traditional photogrammetric feature extraction methods and matching metrics in Photogrammetry (Lingua et al., 2009) is carried out by performing experimental tests on images acquired by Unmanned Aerial Vehicles (UAV) and Mobile Mapping Technologies (MMT) with geometric distortions. Again, the performance of keypoint detectors (FREAK vs. SURF vs. BRISK) are examined in the context of pedestrian detection (Schaeffer, 2013). Işık and Ozkan (2015) evaluated the performance of seven combinations of well-known detectors and descriptors (SIFT, SURF, MSER, BRISK, FREAK, ORB, BRIEF and FAST).

Although many studies have been done on comparison feature detection methods, there is not universally published performance evaluation for stereo image pairs matching of tree. It is necessary to obtain a 3D map of tree to measure geometric characteristics for use in the following items:

- Fertilization: Making a geometric characterization during the productive cycle of trees provides an important part of information required for programming fertilization according to the Nutrient Budgets method (Muhammad et al., 2009; Saa et al., 2013).
- Crop training: Tree canopy is important in order to provide proper training and pruning for trees. The amount of light intercepted by a tree depends on tree density, orientation, size, tree shape and leaf area index (Buba, 2015; Stephan et al., 2008).
- Pest and disease control: The geometric characterization of trees provides the fundamental data that can be used to minimize the environmental impact of pesticides (Russell, 2004).
- Irrigation: Irrigation studies in tree crops are limited by the absence of proper tools for the geometric characterization of vegetation. Therefore, researchers use variables to represent the size and structure of vegetation in some way (Rosell and Sanz, 2012). A precise geometrical characterization of crops at any point during the production cycle may help to establish precise estimations of crop water needs.

In this study, well-known detectors and descriptors in different color spaces were compared. Color spaces are included HSV, H, YCbCr, Y, NTSC and RGB. Twelve combinations of detectors and descriptors which are included: BRISK with SURF (B-F), BRISK with BRISK (B-B), BRISK with FREAK (B-F), Harris with SURF (H-S), Harris with BRISK (H-B), Harris with FREAK (H-F), SURF with SURF (S-S), SURF with BRISK (S-B), SURF with FREAK (S-F), MSER with SURF (M-S), MSER with BRISK (M-B), MSER with FREAK (M-F). For the evaluation, the precision and recall metrics are used by considering the relation between correct matches with number of keypoints in the reference image and the reference image that have been matched.

2. Materials and methods

2.1. Stereo vision system and image acquisition

To calculate tree volume (Immersion method) and to verify results of stereo vision system, an artifact tree (cherry) was made with dimensions of 50 × 70 cm in conical shape. The artifact tree was positioned in imaging room which is equipped with a controlled lighting system and the capability for installing cameras at various heights and angles. The pair of images were acquired using camera (Nikon COOLPIX P510) for both right and left sides. Camera was set to manual mode according to Table 1. Camera was manually placed at two varying viewpoints: with fix height (along y), with 50 cm baseline (along x), with distance of 100 cm from tree (along z) and with parallel optical axes.

2.2. Camera calibration

It is an important pre-step in correctly matching the stereo images and in precisely computing the depth in the stereovision system. It is the process of estimating intrinsic and extrinsic parameters of the camera to minimize the discrepancy between the observed image features and their theoretical positions in the camera pinhole model. Camera calibration in the present study was done based on Heikkila and Silven (1997) and MATLAB camera calibration toolbox (Bouguet, 2004).

2.3. Un-distortion

In theory, it is possible to define a lens that will introduce no distortions. In practice, however, no lens is perfect. This is mainly for reasons of manufacturing; it is much easier to make a spherical lens than to make a more mathematically ideal parabolic lens. It is also difficult to mechanically align the lens and imager exactly. Here we describe the two main lens distortions and how to model them. Radial distortions arise as a result of the shape of lens, whereas tangential distortions arise from the assembly process of the camera as a whole (Bradski and Kaehler, 2008). After obtaining the parameters of the distortion and internal camera calibration, images were undistorted.

2.4. Rectification

For a given point in one image, its correspondent point along an epipolar line in the other image has to be searched (Zhang, 1998). Generally, epipolar lines are not aligned with the coordinate axis and not parallel. Such searches are time consuming since we must compare pixels on skew lines in the image space. These types of algorithm can be simplified and made more efficient if epipolar lines are axis aligned and parallel so that epipolar lines in the original images map to horizontally aligned lines in the transformed images. This can be realized by applying 2D projective transforms to each image. This process is known as image rectification (Loop and Zhang, 1999). Calibrating the stereo system leads to a simpler technique for rectification.