Tree trunk diameter and species estimation using a stereo camera, deep learning and image processing technique

Kenta Itakura, Fumiki Hosoi

Graduate School, University of Tokyo, Tokyo 113-8657, Japan

Abstract

Tree trunk diameter and tree species are two of the most important parameters in investigating trees in urban spaces and forests. The parameters have been conventionally measured and judged by tree caliper and our eyes, which is tedious and time-consuming. It would be preferable if a method based on deep learning techniques which directly conduct the regression is available. However, a large number of training data will be needed because the tree skin and its background are widely varied. Further, it is difficult to obtain huge amount of the agricultural-relevant data. Therefore, a method which can be done with a few data will be desired. In this study, a stereo camera was used for the image acquisition and the scale of the target was determined. Then, the tree trunk diameter and tree species could be estimated using image processing and depth information, then its GUI was constructed. This work can contribute to labor-less and effective tree investigation.

1. Introduction

Trees in urban spaces and forests play an important role such as absorbing CO_2, NO_2, and oxidants, mitigating urban heat islands, offering biodiversity, and manage water cycles. Then, the accurate measurement of tree location and structure is necessary for the wide variety of applications [1]. Tree trunk diameter and tree species are two of the most important parameters. The parameters help, for example, estimating biomass, forest resources, and evaluation of a disaster. The tree trunk diameter is, conventionally, measured using a tree caliper manually. The tree species is judged by our eyes, which is tedious and time-consuming. Furthermore, this manual measurement sometimes entails a pair of persons for manual measurement and the recording. If this measurement can be done with a hand-held camera while moving and the automatic analysis for retrieving the parameters, it leads to labor-less and effective a tree investigation.

A method based on deep learning techniques which directly conduct the regression might be available. It would be done if a large amount of data can be prepared. This method is useful in terms of the needlessness of defining the features. However, such a great number of data is difficult to be collected especially in agricultural-relevant tasks. Thus, a method with a few data is preferable here. On the other hand, in the image analysis for tree trunk diameter estimation, the edge points of tree trunk in the captured 2D images should be determined. The tree trunk and background have a similar color information, so that it is difficult to estimate the tree trunk diameter directly, however, it can be done with a depth information.

From these reasons, we proposed a method which combines a consumer-based stereo camera, image processing, and deep learning-based technique. In this study, estimation of tree trunk diameter and tree species with a stereo camera was conducted.

2. Materials and Methods

2.1. Plant materials and camera device

In total of 2933 images of tree trunks including 23 species such as Himalayan cedar (*Cedrus deodara*), Japanese zelkova (*Zelkova serrata*), and ginkgo (*Ginkgo biloba L.*) were taken for the tree trunk detection and tree species classification. Those images were taken in Shinjuku Gyoen National Garden, Japan, the campus of the University of Tokyo, and the city streets in Tokyo, Japan. The images were taken while walking with smart phones. About three images per one tree sample were recorded with different angles and distances, so that sufficient number of images were collected. The stereo camera used was ZED stereo camera (Stereolabs, San Francisco, USA). The resolution of right and left camera was 720 × 1280 pixels.

2.2. Tree trunk detection

The overflow of the tree detection was illustrated in Fig. 1. First of all, a tree trunk was detected from the captured left image using YOLO 9000 [2]. For the training, 90 % of the collected images above were used. If a tree trunk is detected, the region corresponding to the trunk was cropped. Next, the tree trunk diameter estimation and tree species estimation were done with the cropped image. The tree trunk diameter estimation follows: the depth information was reconstructed with the left and right image from the stereo camera. Here, camera parameters were estimated beforehand with a checkerboard. Then, a distance per one pixel was calculated from the recovered depth (3D) information. Finally, the diameter was estimated with the edge points as represented using the star mark in Fig.1 and the estimated distance per one pixel.

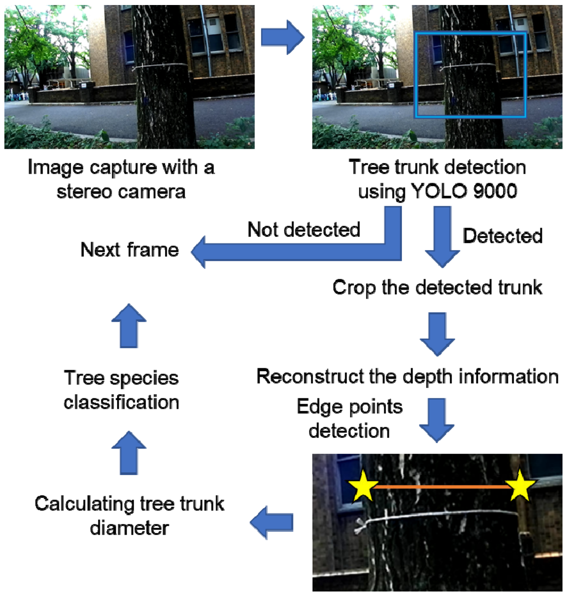

Fig. 1. Workflow for the tree trunk diameter and species estimation.

2.3. Tree species classification

The tree species was classified with support vector machine (SVM) after the calculating the tree trunk diameter. The input variables for the classification were obtained after feature extraction with a pre-trained network, Alexnet [3]. The features with the input tree trunk image before the final fully connected layer were input. The extracted features were trained with the labels of its tree species. Eighty and twenty percent of the collected images were used for the training and test images, respectively.

3. Results and discussion

The tree trunk could be detected with the trained YOLO 9000 accurately. The PR (Precision/Recall) curve was described for evaluating the trained detector performance and its average precision was 0.97, which indicates high performance. Additionally, other samples that were not included in the training/test data set were tested for the tree trunk detection with a GUI enabling to test the sample outdoor while moving as shown in Fig. 2. As a result, almost all of the new samples could be successfully detected. This would be because wide variety of training data with, for example, different light intensity, species, distance and angle toward the target from the camera could be collected. However, a greater number of samples should be tested for validating its feasibility in our future work.

The error of the tree trunk diameter estimation was validated and the RMSE was about 3.14 cm, which would be acceptable for the application. For the tree trunk diameter estimation, a cylinder fitting was reported [4]. However, in this study, the estimation could not be done accurately with the method due to the noises in the reconstructed 3D information from stereo images,

The tree species was classified with the accuracy 90.0 % from the cropped tree trunk image. The amount of test data for the tree species classification at each class was not uniform, then a macro F-measure was also calculated and it indicates the high value of 83.3 %. Generally, the species classification can be done with the tree trunks images to some extent by our sight. However, the classification with a higher accuracy necessitates both tree trunk and leaf images. In the future, a classification system including the trunk and leaf images will be constructed.

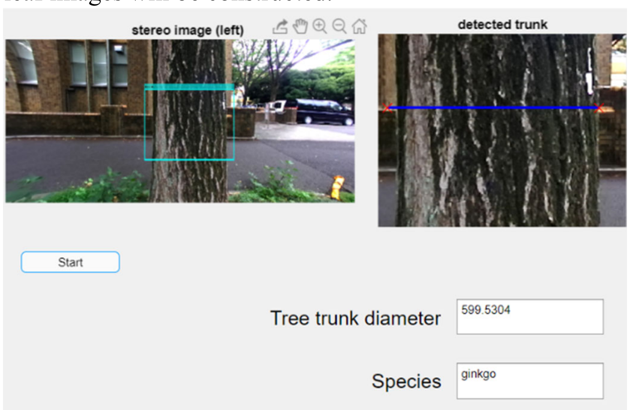

Fig. 2. A GUI for the tree trunk and species estimation. The stereo images were automatically loaded here and the estimation results were shown here.

References

[1] Morgenroth, J., and Ch Gómez. Assessment of tree structure using a 3D image analysis technique—A proof of concept. Urban Forestry & Urban Greening 13 (1), 198-203, 2014.
[2] Redmon, Joseph, and Ali Farhadi. YOLO9000: better, faster, stronger. In Proceedings of the IEEE conference on computer vision and pattern recognition, 7263-7271. 2017.
[3] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems, 1097-1105. 2012.
[4] Fan, Yongxiang, Zhongke Feng, Abdul Mannan, Tauheed Khan, Chaoyong Shen, and Sajjad Saeed. Estimating tree position, diameter at breast height, and tree height in real-time using a mobile phone with RGB-D SLAM. Remote Sensing 10(1), 1845, 2018.