Computers and Electronics in Agriculture 93 (2013) 121–128

Contents lists available at SciVerse ScienceDirect

Computers and Electronics in Agriculture

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

# Tree feature extraction using image data obtained through virtual field server

Xuefeng Wang [a,*], Masayuki Hirafuji [b], Xiaodong Li [a]

[a] The Research Institute of Forest Resource Information Techniques, Chinese Academy of Forestry, CAF, Beijing 100091, China
[b] Field Monitoring Research Team, NARC, Tsukuba 305-8666, Japan

ARTICLE INFO

*Article history:*
Received 6 December 2011
Received in revised form 6 February 2013
Accepted 17 February 2013

*Keywords:*
Virtual field server
Computer vision
Feature extraction
Image understanding
*Xanthoceras sorbifolia*

ABSTRACT

The application of field servers is proving to be increasingly crucial to the process of remote monitoring. These devices are built to continuously obtain large amounts of environmental and meteorological data and, at the same time, transmit back a vast quantity of in situ imagery. The question of how to more effectively utilize these data must be answered. This paper discusses the reconstruction of spatial information, as well as the collection of this information through technical methods. These actions are performed using computer vision based on field server imagery. In order to test and verify the technical approaches involved, such as calibration, matching, reconstruction, and so forth, images of the *Xanthoceras sorbifolia* Bunge tree were used. Two samples of *X. sorbifolia* seedling imagery were reconstructed. It was determined that the precision of the above results was satisfactory. These results demonstrate that the technical approaches can further extract deep information from images obtained through virtual field server. The calculation of image feature points for regular objects, in combination with affine geometry theory, can effectively shield image noise and lead to satisfactory results. Using the sum of the least squares dispersion, in combination with the epipolar line, one can reduce occurrences of image complexity (image matching that occurs during image reconstruction).



## 1. Introduction

In recent years, integrated measurement methods using various different sensors ($CO_2$, soil moisture, photon, and so on) were used more and more. These methods includes control engine, data transmission, set of sensors, power system and other parts, and can automatically measure and receive remotely controlled signals. We call such equipment "Field Server".

Considering the scarcity of resources on earth, finding a sustainable energy alternative has become increasingly important for both stakeholders and the public at large. Biofuel is regarded as one solution. *Xanthoceras sorbifolia* Bunge, which grows in the northern region of China, is an ideal biofuel resource. This tree's seeds contain high rates of oil and sulfate, and it is nitrogen free. However, existing knowledge of this species remains inadequate. This lack of data has stalled the use of this tree as a mass produced biofuel alternative. The lack of available knowledge largely results from the semi-wild pattern in which *Xanthoceras* grows, which makes it difficult for researchers to obtain useful growing data.

Field servers, which have a capacity to sample plots and transfer data via the Internet, act as integrators between agriculture-forestry networks and provide a new means of monitoring *Xanthoceras* (Wang et al., 2006). Currently, the greatest obstacle in the expansion of field server usage is the efficiency and effectiveness of using the substantial amount of data that they collect, especially in relation to image reconstruction.

Most field server is equipped with camera, and some even are equipped with more cameras. These cameras, the same as other sensors, capture live images at intervals set by user. But, these field servers just send back some live images and we view monitoring sites through the images but can do nothing else about the images so far. From the perspective of transmission cost and storage space, it is a waste to use field server. As these images are sent back from fixed camera, and are high density according to timeline, there is advantage to use these images, from the perspective of image analysis. It is much more useful and practical, if further data information can be mined from these images, compared with just viewing images.

The development and application of image data could have a broad range of possible uses, especially considering current conditions (Wu et al., 2004). These uses could include the sorting of fish, the detection of apple surface area, disease analysis, and so on (Yang, 1994; Zion et al., 1999; Vízhányó and Felfoldi, 2000). The objective of this paper is to briefly discuss the technical approaches used to

* Corresponding author. Tel.: +86 10 62889391; fax: +86 10 62888315.
*E-mail addresses:* xuefeng@caf.ac.cn (X. Wang), hirafuji@affrc.go.jp (M. Hirafuji), lxd@caf.ac.cn (X. Li).

0168-1699/$ - see front matter 
http://dx.doi.org/10.1016/j.compag.2013.02.006

extract tree feature using remote plant spatial information based on virtual field server imagery. This discussion can in turn provide a constructive guide for the future utilization of field server imagery.

3D reconstruction is both the primary function of human vision and the primary focus of research areas related to computer vision. These areas include image de-noising, image segmentation, camera calibration, feature point extraction, image matching, and so on. Some related research topics include using multiple cameras to remove adherent noise in the view field (Yamashita et al., 2007), finding edges and lines in imagery (Canny, 1983), the snake method (Kass et al., 1988), and graph cuts (Boykov and Funka-Lea, 2006). Because of length restrictions, this paper will primarily focus on camera calibration, image matching, and reconstruction of crown width.

Camera calibration is a basic and necessary step in 3D reconstruction. The importance of this step is due to its close association with the precision of 3D reconstruction. Camera calibration techniques can be roughly classified into two categories: photogrammetric calibration (Tsai, 1986; Faugeras and Toscani, 1987) and self-calibration (Manbank and Faugeras, 1992). The camera calibration technique reported by Zhang (2000) was adopted for this paper.

Image matching also plays an important role and is applied broadly within various fields. Image matching has been a fundamental technique in associated research projects. It has also undergone a long research history. Two examples of this history are the dynamic programming techniques for matching (Baker and Binford, 1981) and multi-scale matching (Hannah, 1989). As evidenced by Jiang et al. (2004) and Wang et al. (2005), among others, the research on image matching is related to abundant literature on 3D reconstruction algorithms. This paper only examined the effects of radiometric distortions. It did not consider the effect of geometry distortion on image matching because, in laboratory tests, photographic images were only affected by illumination, the direction of the radiating face to photographic objects, and the lens decay of the cameras. The sum of the least squares dispersion method was applied for image matching. It is not difficult to reconstruct 3D models in order to match points for viewing from geometric angles. Thus, the least squares method, which is based on the pin-hole model, was adopted in reconstructing seedling crown width.

This article extracted crown width characteristics of *Sorbifolia* based on images obtained through virtual field server, containing camera calibration, matching and reconstruction. We investigated all the details of the specific algorithm and proposed the corresponding solutions since every part of error will cause the failure of the reconstruction of crown width. Fast matching and tracking algorithm were proposed suitable for specific applications, as well as more intuitive reconstruction equation.

## 2. Material source and reconstruction route

### 2.1. Materials and experimental apparatus

We use virtual field server to monitor indoor seedling growth. The total dimension of the experimental arena was 1800 mm × 1500 mm × 400 mm (Fig. 1). The area was partitioned into 30 smaller areas, each 300 × 300 mm. *X. sorbifolia* seedlings were then planted, and each seedling was numbered with a unique code. A special water vat exhibiting good water permeability was positioned in the left corner of the experimental area so that water could penetrate through the vat and into the soil at a slow rate. This positioning allowed soil closer to the vat to encounter a greater rate of soil moisture. Two groups of different light sources were positioned above the test area (Fig. 1). They were set to turn on and off automatically according to sunset and sunrise. Moreover, light intensity could be adjusted according to time variation. The left light source was positioned near ground level to augment the light intensity, and the right

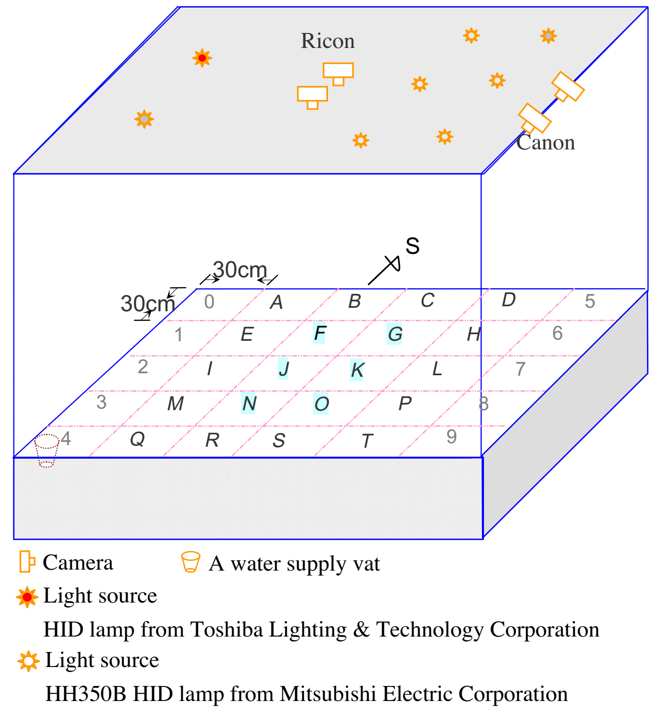


**Fig. 1.** The test area and experimental apparatus and instrumentation.

light source was positioned far from ground level to lessen the light intensity. Two Canon PowerShot S21S cameras with a resolution of 2592 × 1944 were positioned above the test area to view the 20 plants (from A to T). All the cameras were connected to the virtual field server engine, which automatically recorded soil volumetric water content and captured photographs at 10 min intervals.

When the seedlings sprouted, the height and crown width of each seedling was measured in an east–west and south–north direction every morning at 8:30 am. The average measurement taken from both directions was ascertained and then designated as the mean seedling crown width. The experiment lasted for 107 days in total.

### 2.2. Camera model

Several models can be used for camera calibration. The most standard and widely used are the pinhole-based camera models (Forsyth and Ponce, 2008; Deguchi, 2000). If $\tilde{u}$ and $\widetilde{\mathbf{X}}$ are homogeneous coordinates of image points $(u, v)$ and their corresponding spatial points are $(x_1, x_2, x_3)$, relationships between the two can be represented using pinhole camera models. The relationships are found by using the following formula (Roger and Tsai, 1987):

$$\tau\tilde{u} = \mathbf{K}(\mathbf{R}|-\mathbf{R}t)\widetilde{\mathbf{X}} \tag{1}$$

where $\tau$ is a non-zero constant; $\mathbf{K}$ is the intrinsic parameter matrix of the camera; $\mathbf{R}$ is the rotation matrix; and $t$ is the translation vector. The formula unfolds as follows:

$$\mathbf{K} = \begin{pmatrix} a & c & u_o \\ 0 & b & v_o \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{R} = (r_1 r_2 r_3) = \begin{pmatrix} r_{00} & r_{01} & r_{02} \\ r_{10} & r_{11} & r_{12} \\ r_{20} & r_{21} & r_{22} \end{pmatrix},$$

$$\mathbf{t} = \begin{pmatrix} t_0 \\ t_1 \\ t_2 \end{pmatrix}, \quad \tilde{\mathbf{u}} = \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}, \quad \widetilde{\mathbf{X}} = \begin{pmatrix} \mathbf{X} \\ 1 \end{pmatrix} = \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ 1 \end{pmatrix}$$

where $a$, $b$, $c$, $u_o$, and $v_o$ are the internal parameters of the camera (relative to the physical camera); $r$ and $t$ are the external