Rape Plant NDVI Spatial Distribution Model based on 3D Reconstruction

Yang Chen, Yong He*

College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou 310058, China

A paper from the Proceedings of the
14th International Conference on Precision Agriculture
June 24 – June 27, 2018
Montreal, Quebec, Canada

Abstract. *Plants' morphology changes in their growing process. The 3D reconstruction of plant is of great significance for studying the impacts of plant morphology on biomass estimation, illness and insect infestation, genetic expression, etc. At present, the 3D point cloud reconstructed through 3D reconstruction mainly includes the morphology, color and other features of the plant, but cannot reflect the change in spatial 3D distribution of organic matters caused by the nutritional status (e.g. chlorophyll content), as well as illness and insect infestation of the plant. Multispectral photography can reflect the distribution of organic content and other chemical values, and has been extensively applied in such fields as near-ground remote sensing and non-destructive test of quality of agricultural products. In this paper, 31 pieces of multispectral images of 4-leaf rapes are collected for spatial 3D reconstruction through the structure from motion (SFM) method to obtain the 3D point clouds of the rape and filter noise points therein. The maximum length deviation is found to be 0.1023cm and the RMSE is 0.052599 after the model obtained is evaluated with the control point and length, indicating that this method is of good spatial uniformity and accuracy in reconstructing the model obtained. NDVI exponential spatial distribution is then calculated. It is finally proved that the model obtained is of great significance for studying the spatial distribution of plant nutrition, illness and insect infestation in the future.*

Keywords. *morphology; images; rape; 3D reconstructing; structure from motion; NDVI spatial distribution*
List both specific and general terms that will aid in searches.

1 Introduction

Plant morphology is an important basis for study on species identification, vegetative growth, illness and insect infestation, and it is of great significance to accurately and rapidly obtain the plant's 3D morphology for plant phenotypes. Instruments [1] extensively applied for 3D structure reconstruction of plant include Fastscan, Vivid9i, Inspeck, etc. that can achieve relatively accurate 3D surfaces. However, all the results obtained only include spatial structure information, but exclude color, texture and other details. Moreover, it cannot effectively solve the problems regarding 3D structures at different spatial levels due to overlapped leaves, and raises relatively high requirements for the environment. 3D reconstructing by non-contacting images makes it possible to obtain accurate spatial 3D structure of plant and bring the point cloud data obtained colors (RGB). However, the effect of 3D reconstructing may be easily disturbed by ambient light, thus resulting in separated 3D point clouds.

The near infrared reflectance spectroscopy (NIRS) of plant leaf contains lots of information. Moron et al [2] determined the Nitrogen of fresh wheat leaf with NIRS, modelled by PLS, and conducted cross validation by leave-one-out method. Among them, near-infrared band and visible band witnessed the best modelling effect (R2=0.89, SECV=0.64%). Feng Lei et al [3] classified the rice leaf blast applying the multispectral imaging technology, and research shows that the recognition rates for rice seedling blast during vegetative growth phase and rice leaf blast are 98% and 90% respectively. In illness and insect infestation diagnosis based on imaging spectroscopy technology, the crop's growth stage can significantly affect the diagnosis effect. It can be seen that nutritional status, illness and insect infestation, and other physiological and biochemical information of the plant can be obtained from the reflection maps in near infrared region.

In this study, multispectral images of the rape are applied for 3D reconstructing with a view to evaluate the model spatial distribution and calculate and visualize NDVI exponential spatial distribution. It is a non-contacting, easy-to-operate and transferable method that can be used for obtaining accurate 3D model without any large instruments and supporting devices. Firstly, a total of 31 pieces of multispectral images are collected from multiple angles, which will be then performed with blank correction and distortion correction. Then, structure from motion (SFM) algorithm is applied for 3D reconstructing of multispectral images to obtain 3D point cloud data. The statistical filtering is then performed for point clouds with the spatial uniformity of 3D point clouds evaluated. Finally, the spatial 3D NDVI distribution is calculated and visualized.

2 Collection and Pretreatment of Multispectral Images

Plant morphology is an important basis for study on species identification, vegetative growth, illness and insect infestation, and it is of great significance to accurately and rapidly obtain the plant's 3D morphology for plant phenotypes. Instruments [1] extensively applied for 3D structure reconstruction of plant include Fastscan, Vivid9i, Inspeck, etc. that can achieve relatively accurate 3D surfaces. However, all the results obtained only include spatial structure information, but exclude color, texture and other details. Moreover, it cannot effectively solve the problems regarding 3D structures at different spatial levels due to overlapped leaves, and raises relatively high requirements for the environment. 3D reconstructing by non-contacting images makes it possible to obtain accurate spatial 3D structure of plant and bring the point cloud data obtained colors (RGB). However, the effect of 3D reconstructing may be easily disturbed by ambient light, thus resulting in separated 3D point clouds.

2.1 Instruments and Apparatuses

In this study, Tetracam ADC broadband multispectral camera is applied to collect multispectral images from three near infrared channels of red, green and large than 920nm, see Table 1 for

specific parameters. Being portable, this camera has a size smaller than 3CCD multispectral cameras like MS4100, and can collect images without access to the computer. With excellent spectrum curve QE (quantum efficiency), it is suitable for low-altitude remote sensing operations in the agriculture sector [4].

Table 1 Tetracam ADC parameters

Parameter indexes	Value
Size/ mm	122×78×45
Weight/ g	640
Pixel/ 10^6pixel	3.2 (2 048× 1 536)
Sensor size/ mm	6.59×4.9
Spectrum	Green, Red and Near Infrared
Lens	8.5mm standard C-mount lens
Data interface	USB1.1 output, GPS serial input, Trigger output
Storage	Built-in 2GB CF card
Output format	DCM10bit (3M each)
Exposure/ ms	15

2.2 Multi-view Image Collection

This test is done in the evening to avoid the impact of sunlight on images collected. Rapes used for test are 4-leaf winter rapes planted in greenhouse, which are all healthy with 4 leaves. Two 150W sunlight lamps with the color temperature of 6000K that is near to the color temperature of sun in the noon are used for offering continuous light for the visible and infrared spectrum areas. Approximate parallel lights are generated through concave mirror and irradiate on the test bench in an angle of 45°.

The Teflon whiteboard images under test light environment are collected before image collection for blank correction. A total of 31 images are taken around the plant from 3 angles (0°, 30° and 60° with the horizontal plane) with the test bench plane, 12 images in 0° and 30°, and 7 images in 60° respectively. Photography holder is applied for angle control. All the measures are aimed to ensure reasonable overlapping regions between adjacent images, thus offering more feature points for multi-view 3D reconstructing. Checkboard plate is laid under the rape plant to enrich the background details, improve the success rate of image matching and 3D reconstructing, and validate the uniformity and accuracy of the reconstructed scene. See Fig. 2a for the test conditions and Fig. 2b for images obtained.

2.3 Blank Correction and Camera Calibration

Effect of camera calibration significantly affects the accuracy of 3D reconstructing. Camera calibration is divided into image blank calibration and distortion model computing. In blank calibration, Teflon plate of the camera is used for calibrating with PixelWrench. Distortion correction is aimed to calculate the camera's parameters to reduce errors in image matching and spatial positioning. The pinhole camera model is applied for lens calibration to shoot 6 images of checkerboard under different angles, with the checkerboard coverage under each angle of over 60%. The whole checkerboard must be covered. The camera parameters for images shot must be calibrated and performed with distortion correction in Matlab2014b.

a. Experiment conditions

b. Original rape plant images

Fig 1. Experiment conditions and original images

3 Statistical Filtering for 3D Reconstructing and Point Clouds

3.1 Algorithm Principle of 3D Reconstructing

As for structure from motion (SFM), it is unnecessary to know the exact position of the camera during photography or to calibrate with structured light. It only needs to ensure that the object is static 3D object and that the images shot cover the entire object. Motion estimation is firstly made considering the geometrical relationship of 2D images in multi angles, and bundle adjustment is applied to optimize the motion calculation and determine the 3D space, thus extending the point clouds to obtain dense 3D point clouds.

3.1.1 Pinhole Model

The pinhole model, also known as linear model, is extensively applied in the field of computer vision. The multispectral camera applied in this experiment can also be deemed as a pinhole model. The relation of spatial 3D point and image's 2D point is as follows:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} (1)$$

$$\mathrm{K} = \begin{bmatrix} \alpha_u & s & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} (2)$$

In Formula (4), $\widetilde{\mathrm{X}} = [X \quad Y \quad Z \quad 1]^T$ refers to the coordinates of spatial point of the physical object in the world coordinate system, $\tilde{u} = [u \quad v \quad 1]$ refers to the coordinates of points in the camera coordinate system, R is the 3*3 camera orientation matrix, and T refers to the transformation of camera coordinate systems. Under the pinhole model, Formula (1) refers to the position of the object on the image after translating from 3D world coordinate system and rotating. P is the projection matrix. For common camera, it is deemed that each light sensing unit is foursquare, thus the transfer scale $\alpha_u = \alpha_v = \mathrm{f}$, i.e. focal length of the camera. The gradient s is 0, and the principal point is the coordinates of the crossing point of optical axis and imaging plane. Under 2D view, motion refers to the 3D structure of the object obtained from its translation (T) and rotation (R).

Pinhole camera will compute the camera's outer parameters (R and T) when estimates interframe motion. Assuming that X′ is the coordinates in camera reference coordinate system C′, the position of point X in camera coordinate system C is as follows:

$$\mathrm{X} = \mathrm{RX}' + \mathrm{T} (3)$$

Multiplying $X^T[T]$ on both sides:

$$X^T[T]RX' = 0 (4)$$

Given that $\mathrm{E} = [T]R$, the correlation between point X in camera coordinate system and pixel location point u is:

$$u^T(K^{-1T}EK'^{-1})u' = 0 (5)$$

Given that $\mathrm{F} = K^{-1T}EK'^{-1}$, it is a basis matrix. F matrix is a 3*3 matrix with the order of 2, and relates to camera's internal parameters, R and T only. F contains 9 unknown parameters:

$$\mathrm{F} = \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} (6)$$

Assuming that there are n pairs of $u_i = [u_i \quad v_i \quad 1]$ and corresponding points $u'_i = [u'_i \quad v'_i \quad 1]$ between images photographed by two cameras, for the linear equations

$$u^T F u' = 0 (7)$$

There shall be at least 9 points to solve the basis matrix, thus the motion estimation result can be preliminarily obtained.

3.1.2 Bundle Adjustment for Position Optimization

The preliminary motion matrix R and T estimation obtained through motion estimation are insufficient for 3D reconstructing, because there are noises in images and image points in images cannot accurately project corresponding 3D spatial points through the preliminarily estimated camera transfer matrix. Thus, we need to minimize the project errors between the observed and estimated image points applying the bundle adjustment method. In this process, we assume that the characteristic errors of images conform to Gaussian distribution, and perform maximum likelihood search for cost function applying iterative nonlinear least square method. Iterative optimization algorithm method of Levenberg-Marquardt is adopted, which is easy-to-realize and suitable for rapid convergence of preliminary estimation in large scale. The following formula can be used for representing the bundle adjustment process:

$$\min_{a_j b_i} \sum_{i=0}^{n} \sum_{j=0}^{m} v_{ij} d(Q(a_j, b_i), x_{ij})^2 (8)$$

Assuming that there are n 3D points, m view angles, and x_{ij} is the actual projection coordinates of the *i* th point on image j, if point i maps on image j, $v_{ij} = 1; else\ v_{ij} = 0$; $Q(a_j, b_i)$ is the estimated projection coordinates of the *i* th point on image j. This formula is to minimize the project errors of n points on two images.

3.1.3 Generation of Dense Point Clouds

Sparse point clouds obtained by bundle adjustment can sketch out the general structure of the object without any details due to large point distance. Thus, it is necessary to expand the sparse point clouds to obtain more surface information of the object. After the camera projection matrix is obtained, multi-view stereo MVS technology is generally adopted for reconstructing the dense point clouds. In MVS, a point spread-based reconstructing approach is applied to obtain the dense or quasi-dense point clouds of the object surface by matching the sparse spatial points for spreading. This algorithm is suitable for the reconstructing of large scenes, as it is highly flexible and can be achieved without any initial information needed, such as initial bounding box, initial surface, initial disparity range, etc.

3.2 Statistical Filtering

In obtaining point cloud data, there are inevitably noise points in point cloud data due to impacts of instrument accuracy and environmental factors, change in ray-casting and plant surface reflective characters, as well as errors in generation of point clouds through image

matching. The noise points are relatively far from the point clouds. Statistical filtering is achieved based on the distance distribution of point cloud input points and adjacent points, with the points not conforming to error limits removed. The distances between each point and all its adjacent points are calculated, and the results obtained are assumed to be in Gaussian distribution, with its shape determined by the average value and standard deviation. Points with the average distance beyond the standard range (defined by global average distance and variance) are defined as outliers, and removed from the point cloud data.

3.3 3D Reconstructing and Evaluation

The open source program VisualSFM is adopted for 3D reconstructing, operating conditions for which are as follows: Windows7 64bit, Intel I3-3220 3.3 GHz and 8GRAM. Fig. 2 shows the computing of SIFT's key points' location and matching results.

After bundle adjustment, the relation between multi-view optical center position and view angle direction is obtained, and 3,345 sparse point clouds with color information are obtained within 17s.

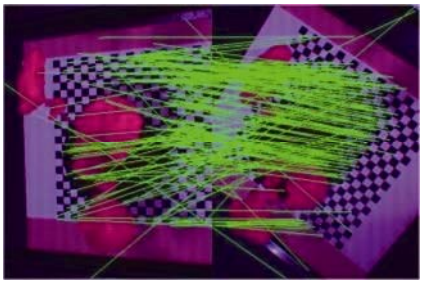

Fig 2. Image key points' location and matching

Sparse point clouds are extended as dense point clouds, which takes 21 min 47 s, to finally obtain the dense point cloud contains 120,089 points, as shown in Fig. 3. According to the image, it is found that 3D point clouds show the plant's morphological characteristics, and the plant's 4 pieces of leaves and upper surface of the flowerpot are properly shown. In the meantime, there are also some drawbacks: 1. There are great number of noise points in point clouds, as some images with missing local key points are filtered in bundle adjustment as large error during matching and reconstructing, such as the 12th, 23th and 32nd images with about 2,000 key points (1,919, 1,916 and 1,963 respectively) that are relatively less than other images; 2. The flowerpot does not miss in point clouds, but only an outline appears. The flowerpot is relatively dark in all images, and its plastic material results in anisotropy, making it is difficult to accurately identify and match the key points; 3. The point cloud contains few 3D points (120,089 points only), thus there are some holes on images.

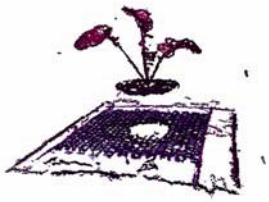

Fig 3. Results of 3D reconstructing

Point Cloud Library (PCL) 1.6.0 open source 3D point cloud processing library is adopted for 3D point cloud statistical filtering. The mean group point number of the filter is set as 50 (i.e. 50 adjacent points are analyzed) and the standard deviation threshold is set as 1.0. The filtering result is as follows: outliers 2,682 and inliers 117,407. The dense point cloud clearly shows the outline of checkerboard paper packed. . The checkerboard is laser printed with the same side length and height, and is laid and closely adhered on the surface of hard test bench to ensure strict uniform scene for reconstructing. According to statistics, the maximum relative deviation,

deviation and root-mean-square deviation are 2.495%, 0.1023 cm and 0.052599 cm respectively.

4 Spatial Distribution of Normalized Difference Vegetation Index (NDVI)

Normalized Difference Vegetation Index (NDVI) is generally used in surface vegetation remote sensing. NDVI is a kind of vegetation index calculated by utilizing green plant's spectral features of low reflectivity to red light ad high reflectivity to near-infrared light, and can better show the growth status of plants. NDVI can be used to predict the plant's growth status, illness and insect infestation, physiological and biochemical information, thus offering theoretical basis for production prediction. Thus, it is of great significance to research the plant's NDVI spatial distribution. NDVI is defined as follows:

$$NDVI = \frac{(NIR-R)}{(NIR+R)} \quad (9)$$

In point cloud file .ply, each voxel contains the color value information of the reconstructed image. For ADC camera, channels 1, 2 and 3 are NIR channel, Red channel and Green channel respectively. NDVI value of current point can be calculated on voxel. The point cloud data is then visualized. Fig. 6a shows the results, and Fig. 6b shows the histogram of NDVI value distribution.

Conclusion

In this study, SFM algorithm is adopted for 3D reconstructing based on multispectral images collected by ADC multispectral camera. The statistical filtering is then performed for point cloud data obtained, which demonstrates that the reconstructed model is of good uniformity and availability, with the deviation controlled within 0.1023 cm. In the meantime, the spatial distribution of the commonly-used vegetation index NDVI is calculated and visualized, thus laying foundation for studies on the plants' spatial nutrient and organic distribution. Compared with laser scanning 3D reconstructing, SFM 3D reconstructing is cost-efficient and no special equipment is needed. It is easy-to-carry and suitable for field operations. The model obtained not only contains structure information but can also introduce the infrared information in the reconstructed model for study on spatial distribution of NDVI.

References

[1] Eder Maximilian, Brockmann Gernot, Zimmermann Alexander, et al (2013). Evaluation of Precision and Accuracy Assessment of Different 3-D Surface Imaging Systems for Biomedical Purposes. Journal of Digital Imaging, 26(2), 163 - 172.

[2] Moron Alejandro, Garcia Adriana, Sawchik Jorge, et al (2007). Preliminary study on the use of near-infrared reflectance spectroscopy to assess nitrogen content of undried wheat plants. Journal of the science of food and agriculture, 2007, 87(1), 147 - 152.

[3] Feng Lei, Chai Rongyao, Sun Guangming(2009). Identification and Classification of Rice Leaf Blast Based on Multi-Spectral Imaging Sensor. Spectroscopy and Spectral Analysis, 29(10), 2730 - 2733.

[4] Tian Zhenkun, Fu Yingying, Liu Suhong, et al (2007). Rapid crops classification based on UAV low-altitude remote sensing. Transactions of the Chinese Society of Agricultural Engineering, 109-116.