

# Plant Three-dimensional Information Detection Based on Single Moving Camera

Dongbo Zhang, Xi Deng, Xiaojing Liu, Bin Zhang

College of Engineering, China Agricultural University, Beijing 100083, China  
zdb0204@cau.edu.cn, zhangbin64@cau.edu.cn

**Abstract** - Using non-contact sensing techniques and facilities to gain plant three-dimensional (3D) information can satisfy different kinds of needs in digital and intelligent agricultural production. 3D information could be used for vegetative modeling and monitoring or plant production management and operation. A plant 3D information detecting technology based on the movement of single camera was proposed. The method was divided into three steps. Firstly, calculate the camera matrix in different spatial positions. Presume that world coordinate was coincided with camera coordinate in one position. With the known movement parameters of camera, it was able to calculate the camera matrix in another position. Secondly, extract plant information using image processing technology. Then, sample plant areas and pick up matching points. A region matching method based on epipolar constraint was applied to obtain corresponding points between images gathered in different positions. The region matching error metrics was the sum of absolute differences (SAD) of RGB channels. Finally, with the matching points and projective geometry, target plant's 3D information could be calculated. A web camera common on the market was used in experiments to verify this detecting technique. It turned out that while the average distance between camera and plant was less than 1.2m, the detecting error of x and y axis was within  $-18.9\text{mm}\sim+16.2\text{mm}$  and  $+7.0\sim+28.1\text{mm}$  in z direction. While the average distance about 1.5m, the detecting error of x and y axis was within  $-36.0\text{mm}\sim+5.6\text{mm}$  and  $+7.0\sim+59.0\text{mm}$  in z direction. And it also appeared that the closer the distance was the higher the detecting precision would be.

**Index Terms** - information technology, sensing detection, three-dimension (3D), plant

## 1. Introduction

A three dimensional (3D) structure of a plant is able to reflect its growth condition and morphologic changes intuitively. Therefore, the 3D measurement and analysis is significant to the understanding and diagnosis of the plant stress responses. The growth condition and morphologic changes of high growth rate plants like crops and vegetables will respond sensitively on different stresses. So, different operating strategies should be applied catering to target plant's growth condition and morphology. Using non-contact sensing techniques and facilities to gain plant 3D information can satisfy different kinds of needs in digital and intelligent agricultural production. 3D information could be used for vegetative modeling and monitoring or plant production management and operation.

In the aspect of getting plant 3D information, researchers found out many different methods. Q. U. Zaman [1] realized fruiter canopy's 3D reconstruction and volume measurement with ultrasonic sensor. Joan R. Rosell [2] applied a 2D laser scanning method to identify fruiter's geometrical parameters

and reconstruct the canopy. Ulrich Weiss [3] researched on plant detection and 3D drawing with 3D laser radar sensor while applying robots in agriculture. Chen Bingqi [4] researched on the measurement of field corn's growth parameters and 3D virtual display of its growing process based on binocular stereo vision. Bernhard Biskup [5] designed a binocular stereo system that was able to detect inclining angle of leaf. A. D. Nakarmi [6] designed a detecting system which can cater to corns in different growing period with TOF camera. Yann Chéné [7] used Kinect RGB-Depth Camera invented by Microsoft to detect the 3D information of a whole plant. Fang Hui [8] researched on methods to attain plant 3D information with a 3D information collection system using structural light.

Machine vision is considered to be a potential, low failure and risk and contactless measuring technology which makes it suitable for being applied to agricultural circumstances. In many agricultural annual rings, sensor and plant move relatively. So, integrated with epipolar and projective geometry, a plant 3D information detecting technology based on the movement of single camera was proposed.

## 2. 3D Information Detecting Algorithm for Single Camera

### A. Epipolar geometry

According to camera's multi-view geometric principle, the imaging point of one point between two different views should meet the epipolar constraint condition which shown in Fig. 1.

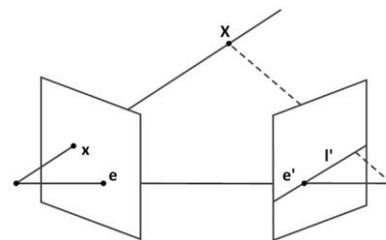


Fig.1 Epipolar geometry

Epipolar geometry is an intrinsic projective geometry between two views in different angle. It is independent from scene structure and only related to inner parameters and relative poses of camera. This kind of intrinsic epipolar geometric relationship could be presented by a  $3 \times 3$  fundamental matrix  $F$  with a rank of 2. Presume that  $x$  and  $x'$  are image coordinates of one point in different views. Then, the image point and epipolar should meet the corresponding relationship that follows.

$$\mathbf{X}^{\text{T}} \mathbf{F}_{\mathbf{X}} \quad (1)$$

$$\mathbf{I}' \mathbf{F}_{\mathbf{X}} \quad (2)$$

The fundamental matrix can be calculated directly by algebraic method with camera matrix. Suppose that world coordinate coincides with one of the images' camera coordinate. The movement vector of camera is  $t$ . So the two camera matrix should be assumed to be:

$$\mathbf{P} = \mathbf{K}[\mathbf{I} \mid \mathbf{0}] \quad \mathbf{P}' = \mathbf{K}'[\mathbf{I}' \mid t]$$

Then  $\mathbf{P}$ 's violation and the centre of camera is:

$$\mathbf{P}^* = \begin{bmatrix} \mathbf{K}^{-1} \\ \mathbf{0}^{\text{T}} \end{bmatrix} \quad \mathbf{C} = \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix}$$

According to reference [9],

$$\mathbf{F} = [\mathbf{P}' \mathbf{C}]_{\mathbf{X}} \mathbf{P}' \mathbf{P}^* = [\mathbf{K}t]_{\mathbf{X}} = [\mathbf{e}']_{\mathbf{X}} \quad (3)$$

In which  $[\mathbf{x}]_{\mathbf{X}} = \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix}$  and  $\mathbf{e}'$  is the subpoint

of the first camera's centre in the second camera. So,  $\mathbf{e}'$  is called epipole.

### B. Acquisition of corresponding points in images

The detecting operation in agriculture is mostly targeting at plant. Thus, in order to enhance algorithm's execution efficiency, only green areas would be sampled and matched. Image of the first view angle was grayed by Supper Green operator (2G-R-B). Then, the maximum between-cluster variance (Otsu) was used to get the segmentation threshold automatically and binarize the Supper Green image. A morphological algorithm was used for filter. By region matching based on epipolar constraint, it was able to sample the green parts and obtain corresponding points.

In the design of region matching algorithm, error was the combination of absolute error and SAD. The system used color images which output with three channels of RGB. So, in order to improve the matching precision, the sum of absolute differences of each channel was firstly calculated separately. Then, add the three errors together.

$$E_{SAD}(\mathbf{u}) = \sum_i (\Delta I_R + \Delta I_G + \Delta I_B) \quad (4)$$

In which  $\Delta I_{\mathbf{X}} = \sum_i |I_{\mathbf{IX}}(\mathbf{x}_i + \mathbf{u}) - I_{\mathbf{0X}}(\mathbf{x}_i)|$ .

The minimum of  $E_{SAD}$  was picked as the alternative point for matching. And a threshold  $N$  was selected as a limitation. When the minimum  $E_{SAD}$  satisfy the following condition, the right information of sample point's corresponding point was got.

$$E_{SAD}(\mathbf{u})_{MIN} < N \quad (5)$$

The threshold  $N$  should be adjusted according to situation of application. The experiments here chose 30 as the threshold.

### C. Method of calculating corresponding points' space coordinates

When corresponding points from different views were got, the 3D information of image points could be achieved via the projective geometry. There were plenty of ways to calculate the points. The liner triangulation method was applied to calculate points' 3D coordinates directly. Presume  $\mathbf{x} \leftrightarrow \mathbf{x}'$  are the corresponding points. Then the liner expression of  $\mathbf{x} = \mathbf{P}\mathbf{X}$ ,  $\mathbf{x}' = \mathbf{P}'\mathbf{X}$  is  $\mathbf{x} \times (\mathbf{P}\mathbf{X}) = 0$ ,  $\mathbf{x}' \times (\mathbf{P}'\mathbf{X}) = 0$ . So, the equation set could be:

$$\mathbf{A}\mathbf{X} = 0 \quad (6)$$

$$\text{In which } \mathbf{A} = \begin{bmatrix} x p^{3T} - p^{1T} \\ y p^{3T} - p^{2T} \\ \mathbf{x}' p'^{3T} - \mathbf{p}'^{1T} \\ \mathbf{y}' p'^{3T} - \mathbf{p}'^{2T} \end{bmatrix}, \quad \mathbf{p}'^{iT} \text{ represents } \mathbf{P}' \text{'s row.}$$

If the singular value decomposition (SVD) method was applied to equation (4), a solution set which satisfy the constraint  $\|\mathbf{x}\| = 1$  could be calculated. Suppose  $\mathbf{A}$ 's SVD is

$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^{\text{T}}$ , in which  $\mathbf{D}$  is a diagonal matrix.  $\mathbf{D}$ 's main diagonal elements are positive and with a descending order. Then the last column of  $\mathbf{X}$ 's corresponding matrix  $\mathbf{V}$  should be  $\mathbf{X} = (X, Y, Z, T)^{\text{T}} = \mathbf{V}_4$ . So the image points' inhomogeneous coordinates are  $(X/T, Y/T, Z/T)^{\text{T}}$ .

## 3. Experiments and System Composition

### A. Image capture and processing facilities

A USB2.0 web camera Blue Enchantress M2200 was used as the image capturing facility. It was really cheap and could be bought on markets. The sensor of camera was 1/5 inch CMOS plate OV7740. The sensor's dynamic solution was 640\*480 and its maximum frame frequency was 30 frames per second. Moreover, it had the functions of automatic white balance and automatic exposure. The camera had five lenses and could be focused manually.

Image processing software used the C/C++ module in Microsoft Visual Studio 2008 as a development platform. The software was developed combining with C language and OpenCV Computer Vision Function Library. And the software was operated on a Lenovo Soyang E46 computer with the operating system of Windows7 32 bit. CPU was Intel Core i3 M 380, 2.53GHz.

### B. Camera calibration

While the camera's moving, the inner parameters do not change. So the calibration method applied was offline. Using the Matlab Calibration Toolbox to calibrate the chessboard images taken from different angles, the inner parameters of the camera were certified. The inner parameter matrix was:

$$\mathbf{K} = \begin{bmatrix} 750.46 & 0 & 360.18 \\ 0 & 748.44 & 213.04 \\ 0 & 0 & 1 \end{bmatrix} \quad (7)$$

The unit of inner parameters was pixel and two decimal places were reserved. In the matrix, the angles of inclination of X and Y imaging flat were set to be 0°.

C. Experiment of single plant's 3D information detection

The experiment was operated in the laboratory. A single pot plant was picked and detected. Then, the results calculated by the system were compared with the ones measured manually. Firstly, set the first angle of camera and measure plant's 3D information manually, Fig.2.

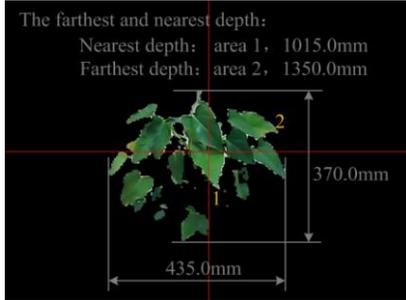


Fig.2 Manually measured values of a single plant

From the first view point, move the camera alongside the camera coordinate system's X and Y axis separately. Both directions needed to be moved positively and negatively by 40mm ( $X \pm 40$ ,  $Y \pm 40$ ). And it also includes a combinative movement of 40mm positively on X-axis ( $X+40$ ) and 40mm negatively on Y-axis ( $Y-40$ ). The detecting results were showed in Table.1.

Table.1 Results of detecting a single plant (mm)

Moving direction	Detecting result	Measuring result	Error
X+40	X: 416.1	435.0	-18.9
	Y: 357.9	370.0	-13.0
	Z: 363.1	335.0	+28.1
X-40	X: 430.6	435.0	-4.4
	Y: 374.5	370.0	+4.5
	Z: 363.1	335.0	+28.1
Y+40	X: 448.1	435.0	+13.1
	Y: 379.5	370.0	+9.5
	Z: 361.0	335.0	+26.0
Y-40	X: 423.2	435.0	-11.8
	Y: 381.1	370.0	+11.1
	Z: 359.0	335.0	+24.0
X+40 Y-40	X: 445.2	435.0	+10.2
	Y: 379.7	370.0	+9.7
	Z: 346.0	335.0	+11.0

According to the results, this way of detection had an error of: -18.9~+13.1 on X, -13.0~+11.1 on Y and +11.1~+28.1 on Z direction. It was obvious that the results on Z direction were entirely higher. That was because the size estimation of depth was directly related to the difference value of pixel between two images. Moreover, the difference value of pixel was a change of integer which made the size change of depth

discontinuous, Fig.3. This kind of non-continuity was typically obvious when the movement was simply alongside X or Y axis. The size of depth was the difference of the maximum and the minimum. While valuing, the maximum was a little higher and, at the same time, the minimum was a little lower. So, the result came out that the detecting sizes of depth were entirely higher.

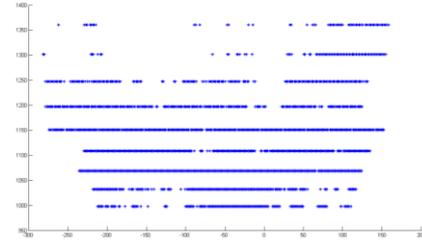


Fig.3 discontinuously changing size of depth

A method of sampling part of a plant point by point was applied to obtain 3D coordinates of the points. And the experiment used a threshold of 30 to eliminate wrong matching points. After achieving the 3D coordinates of the plant in first view angle, a 3D point cloud picture was drawn by Matlab, Fig.4.

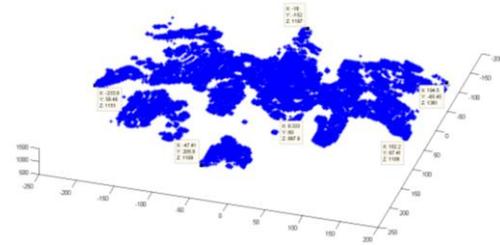


Fig.4 3D point cloud of a single plant

D. 3D information detecting experiment of two plants

In order to testify the adaptation of the algorithm and the affection of different depth of plant, two plants were set at the same time in this experiment. And these plants were located with one nearer and the other farther. Here is the 3D information of plants measured manually, Fig.5.

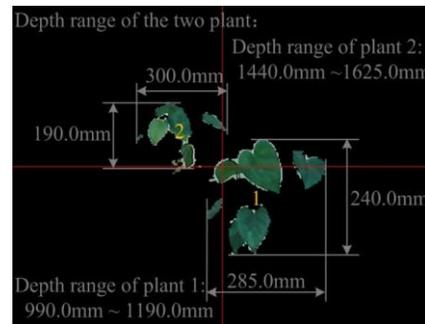


Fig.5 Manually measured values of two plants

Applying the same moving way in the first experiment, the results were shown in Table.2.

Table.2 Results of detecting two plants (mm)

Moving direction	Detecting result of plant 1		error	Detecting result of plant 2		error
X+40	X:	286.2	+1.2	X:	276.2	-23.8
	Y:	241.5	+1.5	Y:	182.9	-7.1
	Z:	215.0	+15.0	Z:	244.0	+59.0
X-40	X:	284.8	-0.2	X:	301.1	+1.1
	Y:	239.3	-0.7	Y:	195.6	+5.6
	Z:	207.0	+7.0	Z:	205.0	+20.0
Y+40	X:	270.7	-14.3	X:	282.8	-17.2
	Y:	256.2	+16.2	Y:	182.7	-7.3
	Z:	214.3	+14.3	Z:	192.0	+7.0
Y-40	X:	294.1	+9.1	X:	264.0	-36.0
	Y:	249.2	+9.2	Y:	188.2	-0.8
	Z:	215.0	+15.0	Z:	237.0	+52.0
X+40 Y-40	X:	276.4	-8.6	X:	293.6	-6.4
	Y:	247.4	+7.4	Y:	194.6	+4.6
	Z:	215.5	+15.5	Z:	214.0	+29.0

The experiment result turned out that plant 1 whose distance from camera was about 1.1m had a detecting error of: -14.3~+9.1 on X, -0.7~+16.2 on Y and +7.0~+15.5 on Z direction. And plant 2 whose distance was about 1.5m from camera had an error of: -36.0~+1.1 on X, -7.3~+5.6 on Y and +7.0~+59.0 on Z direction. What should be noticed was that the main error on X direction in plant 2 was caused by the incomplete segmentation of plant. A 3D point cloud was drawn after getting the points' 3D coordinates, Fig.6.

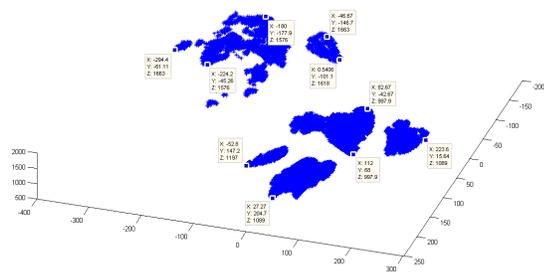


Fig.6 3D point cloud of two plants

#### E. Error analysis

There were three main error sources by using the detecting method of single moving camera.

1. Camera's amount of movement and rotation had certain errors. The rotation error had a bigger effect to the result.
2. The segmentation of plant also had errors. It might mainly cause a larger error between detecting value and measuring value. Moreover, it would at the same time affect data integrity while applied to practical use.
3. The matching error.

By adding sliding pair's structural intensity and stability, the movement error could be reduced. And as the amount of rotation should be avoid or decreased. Furthermore, better segmentation methods could be adopted. In order to improve the matching precision, the sub pixel precision method could possibly raise the precision by an order of magnitudes. Thus,

plant's 3D information detecting precision would be increased by a large margin.

#### 4. Conclusion

A research of plant 3D information detection based on single moving camera was done. By using known moving parameters of camera, the camera matrix in different position could be attained. Then the epipolar constraint relationship between pixels in different images was calculated. The epipolar constraint and the SAD from images' RGB channels together formed an error metric function. This function was used in region matching to obtain corresponding points' information. Then, imaging points' 3D information could be calculated with projective geometry. In order to raise system's working efficiency, only information about plant part was sampled to be matched. So, the Otsu method was adopted to choose threshold automatically and segment the Supper Green image (2G-R-B). The experiments were operated indoor with a web camera. Results showed that when the distance between plant and camera was less than 1.2m, the detecting error of x and y axis was within -18.9mm~+16.2mm and +7.0~+28.1mm in z direction, when the average distance between camera and plant was about 1.5m, the detecting error of x and y axis was within -36.0mm~+5.6mm and +7.0~+59.0mm in z direction. And it also appeared that the closer the distance was the higher the detecting precision would be.

#### Acknowledgment

This work was supported by Research Fund for the Doctoral Program of Higher Education of China under Grant No.20120008110046.

#### References

- [1] Zaman Q.U., Schumann A.W., Hostler H.K. Quantifying sources of error in ultrasonic measurements of citrus orchards. American Society of Agricultural and Biological Engineers, 2007, 23(4): 449-453
- [2] Joan R. Rosell, Jordi Llorens, Ricardo Sanz, et al. Obtaining the three-dimensional structure of tree orchards from remote 2D terrestrial LIDAR scanning. Agricultural and Forest Meteorology, 2009, 149: 1505-1515
- [3] Ulrich Weiss, Peter Biber. Plant detection and mapping for agricultural robots using a 3D LIDAR sensor. Robotics and Autonomous Systems, 2011, 59:265-273
- [4] Chen Bingqi, He Chun, Ma Yanping, et al. 3D image monitoring and modeling for corn plants growth in field condition. Transactions of the CSAE, 2011, 27(Supp.1): 366-372. (in Chinese with English abstract)
- [5] Bernhard Biskup, Hanno Scharr, Ulrich Schurr. A stereo imaging system for measuring structural parameters of plant canopies. Plant, Cell and Environment, 2007, 30:1299-1308.
- [6] A.D. Nakarmi, L. Tang. Automatic inter-plant spacing sensing at early growth stages using a 3D vision sensor. Computers and Electronics in Agriculture, 2012, 82:23-31
- [7] Yann Chéné David Rousseau, Philippe Lucidarme, et al. On the use of depth camera for 3D phenotyping of entire plants. Computers and Electronics in Agriculture, 2012, 82:122-127
- [8] Fang Hui, Hu Lingchao, He Rentao, et al. Research on plant three-dimensional information acquisition method. Transactions of the CSAE, 2012, 28(3): 142-147. (in Chinese with English abstract)
- [9] Richard Hartley, Andrew Zisserman. Multiple view Geometry in Computer vision. 2nd Edition. Cambridge: Cambridge University Press, 2004.