Camera Calibration for 3D Leaf-Image Reconstruction using Singular Value Decomposition

Hermawan Syahputra
Department of Mathematics
State University of Medan, Sumatera Utara, Indonesia

Reza Pulungan
Department of Computer Science and Electronics
Universitas Gadjah Mada, Yogyakarta, Indonesia

***Abstract*—Features of leaves can be more precisely captured using 3D imaging. A 3D leaf image is reconstructed using two 2D images taken using stereo cameras. Reconstructing 3D from 2D images is not straightforward. One of the important steps to improve accuracy is to perform camera calibration correctly. By calibrating camera precisely, it is possible to project measurement of distances in real world to the image plane. To maintain the accuracy of the reconstruction, the camera must also use correct parameter settings. This paper aims at designing a method to calibrate a camera to obtain its parameters and then using the method in the reconstruction of 3D images. Camera calibration is performed using region-based correlation methods. There are several steps necessary to follow. First, the world coordinate and the 2D image coordinate are measured. Extraction of intrinsic and extrinsic camera parameters are then performed using singular value decomposition. Using the available disparity image and the parameters obtained through camera calibration, 3D leaf-image reconstruction can finally be performed. Furthermore, the results of the experimental depth-map reconstruction using the intrinsic parameters of the camera show a rough surface, so that a smoothing process is necessary to improve the depth map.**

***Keywords*—Camera calibration; image reconstruction; 3D leaf images; singular value decomposition**

I. Introduction

Automatically recognizing leaf objects using a computer is a challenging task. The main challenge lies in the image variation. A different position of the camera can see the object in a different variation. The image will also vary depending on the direction and position of the object. To maintain the correct direction and position of the image, a 3D stereo concept can be used. This technique is a stereoscopic technique of computer vision that has been developed by several researchers in the field of agricultural automation [1], [2].

Camera calibration is a necessary step for 3D computer vision to extract information on the distance measurement from a 2D image. It is widely studied in the fields of computer vision and photogrammetry. Developing computer vision algorithms with high accuracy is not easy. Designing computer vision algorithms for two cameras requires knowledge and understanding in how images of the same scene are viewed from different viewpoints. Camera calibration system should be designed in such a way that the image coordinates of the world point will remain detached from the position and direction of the camera.

A calibration process refers to a process of determining intrinsic and extrinsic camera parameters from a number of correspondences between the 3D point and the projection of that point to one or several 2D images [3]. Most frequently, this is done by using the calibration checkerboard or other easily recognized marker patterns [4].

Various algorithms have been presented in the literature to solve this calibration problems. Salvi et al. [5] compared several methods, namely, Halls and Faugeras, who employed the technique of least squares to generate camera parameters, and Tsai and Weng, who used a two-stage technique, where the first stage used a linear approach with the aim of generating initial guesses and the second stage used an iterative algorithm to optimize the parameters [5]. Zhang [3] proposed a calibration method using multiple images of a planar calibration grid.

Many researches also proposed camera calibration for reconstructing 3D objects. Weng et al. [6], for instance, presented a camera model that accounts for major sources of camera distortion, namely, radial decentering, and thin prism distortions. Zhu et al. [7] proposed reconstructing 3D-models of old Beijing city by a structured light photogrammetry.

This paper focuses on calibration problems of a camera with two lenses (stereo camera), where the relative projection matrix between the lenses must be highly accurate. This projection matrix is usually used in the calculation of the depth of the images taken using the stereo camera [8] or for human pose reconstruction [9]. For this application, the quality of the camera calibration has a direct impact on the quality of the overall results.

The method we propose in this paper is based on [10]. Consider Fig. 1, in which an imaging model of two cameras C and C' is depicted.

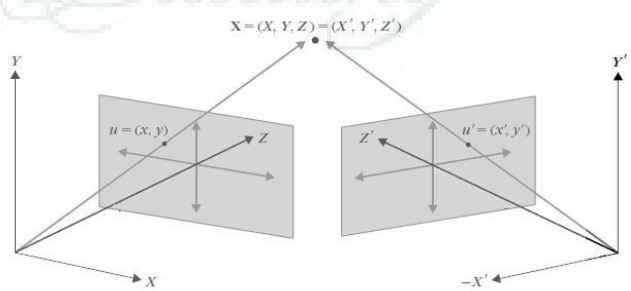

Fig. 1. Imaging an object using two cameras [10].

Fig. 1 shows a model imaging from a camera with double lenses (a stereo camera) or two separate physical cameras

or a camera moving at different positions. Assume that the scene coordinate for a point X in the coordinate system C is (X, Y, Z) and in the coordinate system C' is (X', Y', Z'). Then, $u = (x, y)$ and $u' = (x', y')$ are the coordinates for the image of X that correspond to the planes of the images, P and P', respectively. The points u and u' are called the *corresponding points*.

Therefore, the same scene of the world coordinates is mapped onto a different system of image coordinates. This is to adjust points that are mathematically related by a one-to-one mapping explicitly. In other words, after a camera is calibrated, the scene of the world coordinates is the same and each camera describes the same pixels.

In this paper, 3D leaf images are captured using Fuji Finepix stereo camera Real 3D W3 Camera. This camera has two 10 MP lens, and 3D real-time screen, with no glasses.

The rest of the paper is organized as follows: Section II describes the basis of parameter extraction using singular value decomposition and transformation of coordinate systems. Section III describes the proposed method of camera calibration. Section IV presents the result of our experiments with the proposed method and provides analysis. This chapter also discusses the application of the method on 3D leaf-image reconstruction. Finally, Section V concludes the chapter.

II. Preliminaries

A. Orthogonal and Orthonormal Vectors

Two vectors $\vec{x}$, $\vec{y}$ that are perpendicular to one another are called orthogonal, denoted by $\vec{x} \perp \vec{y}$. In this case, inside/outside multiplication (inner/dot product), denoted by $\langle \vec{x}, \vec{y} \rangle$, is zero. Two vectors are said to be orthonormal if they are orthogonal and have a unit length, *i.e.*, the length (norm) of the vectors equals to one.

B. Singular Value Decomposition

Singular value decomposition (SVD) is one of many factorization methods applied in least-squares calibration to resolve linear equations, to compute the rank and the null spaces of matrices, and so on. A system of linear equations can be written in the form of $\mathbf{A}\vec{x} = \vec{b}$, where $\mathbf{A}$ is a matrix of size $m \times n$, where $m > n$. Then, $\mathbf{A}$ can be factorized using SVD: $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^\top$, where $\mathbf{U}$ is an $(m \times n)$-matrix formed by orthogonal vectors, $\mathbf{S}$ is a diagonal matrix with positive or zero elements of size $n \times n$, and $\mathbf{V}$ is an $(n \times n)$-matrix formed by orthogonal vectors.

The terms above is called the SVD of matrix $\mathbf{A}$. For example, let $\mathbf{S} = diag\{\sigma_1, \sigma_2, \ldots, \sigma_n\}$, where $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$. Therefore, $\sigma_1, \sigma_2, \ldots, \sigma_n$ are regarded as *singular values* of $\mathbf{A}$. Columns of $\mathbf{U}$ and $\mathbf{V}$ are left and right singular vectors, respectively, for the corresponding singular values. For a system of homogeneous linear equations $\mathbf{A}\vec{x} = \vec{0}$, vector $\vec{x}$ in the null space of $\mathbf{A}$ is a solution. Columns of $\mathbf{V}$, whose corresponding singular values approach zero, are solution vectors.

For a system of non-homogeneous linear equations $\mathbf{A}\vec{x} = \vec{b} \neq \vec{0}$, we aim at finding the solution $\vec{x}$ with the shortest length $|\vec{x}|^2$, where $\vec{x} = \mathbf{V}[diag(\sigma_1^{-1}, \sigma_2^{-1}, \ldots, \sigma_n^{-1})]\mathbf{U}^\top\vec{b}$. In this case, for each singular value $\sigma_i = 0$, σ_i^{-1} is replaced by 0. If $\vec{b}$ is not in the range of $\mathbf{A}$, then vector $\vec{x}$ that satisfies the system of equations cannot be found. Moreover, it is impossible to generate an exact solution.

If $n = m$, *i.e.*, $\mathbf{A}$ is a square matrix, SVD is determined by $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^\top$, where $\mathbf{U}$ and $\mathbf{V}$ are orthogonal square matrices. Then, $\mathbf{A}^{-1} = \mathbf{V}[diag(\sigma_1^{-1}, \sigma_2^{-1}, \ldots, \sigma_n^{-1})]\mathbf{U}^\top)$. The range of $\mathbf{A}$ is the same as the rank of $\mathbf{A}$. If $\mathbf{A}$ is singular, the rank of $\mathbf{A}$ is therefore lower than n, where $n = rank(\mathbf{A}) + null(\mathbf{A})$. Columns of $\mathbf{U}$, associated with singular values that are not equal to zero, is an orthonormal set for the range of $\mathbf{A}$. Columns of $\mathbf{V}$, associated with singular values that are not equal to zero, is an orthonormal set for the null space of $\mathbf{A}$.

III. Proposed Method

The proposed method is oriented to the final objective, *i.e.*, to bring together different camera views by ignoring lens distortions. Fig. 2 depicts the proposed calibration method.

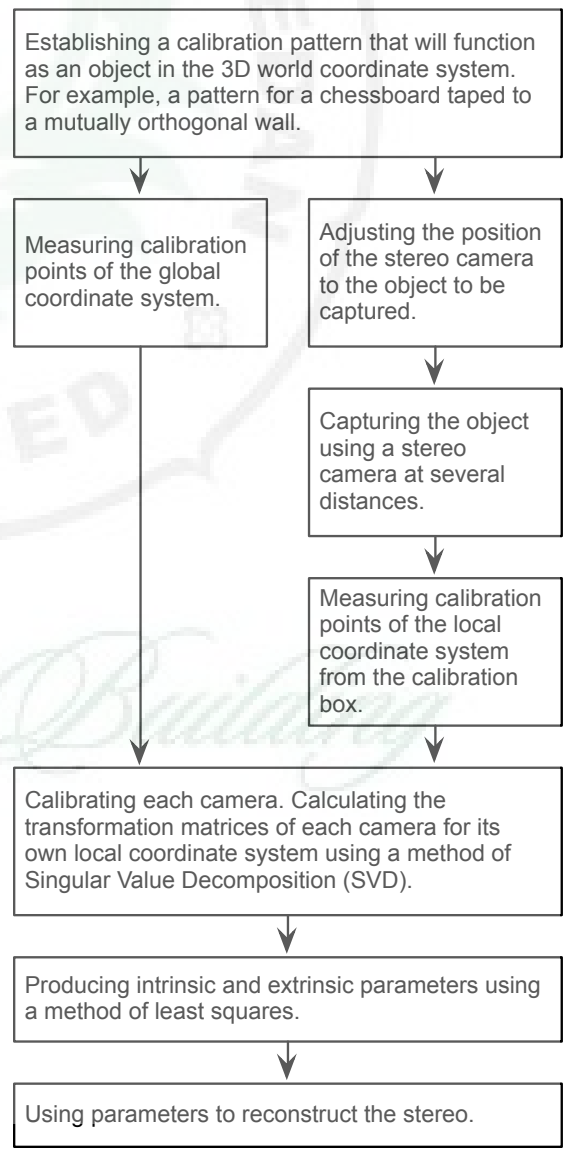

Fig. 2. The proposed calibration method.

At the initial stage, the calibration pattern is built using a checkerboard pattern. The checkerboard uses size of 2 cm for each box and is placed in the wall angled to the real world. Some vertices of each checkerboard box are calculated as 3D coordinates and are used as calibration points (global coordinates). Furthermore, a checkerboard image from the camera portrait is taken and the coordinates of the pixels from the predetermined points are measured as 2D coordinates (local coordinates). Once the camera parameters of the local coordinate system have been calculated, the global coordinate data are then used to transfer all local coordinate data to the same global coordinates to register everything into the same global coordinate system.

After all these operations are done, the image's coordinates are calculated using pixel coordinates of the predetermined calibration points within the space. It can be seen that although the world coordinates of the image remain the same, they are different in terms of the position and direction of the camera. Then, the system is calibrated in such a way that the coordinates of the image from this point are not different from the position and orientation of the camera, or in other words, the whole system is calibrated.

Furthermore, the parameter values generated are subsequently used to reconstruct the stereo images, whose disparity values have been obtained. There are five intrinsic parameters, namely, a scale factor s, an image center (u_0, v_0), and (alpha and beta) focal lengths f. They are not independent. The focal lengths are used to construct depth map.

A. Camera Parameters Extraction

To express a point of any object in the world coordinate system, it is necessary to first transform the point into the camera's coordinates. This transformation consists of translation and rotation. Following [10], let the coordinate of the world is denoted by $P_w(x_w, y_w, z_w)$ while the coordinate of a 3D camera is denoted by $C(x, y, z)$. Furthermore, the transformation of the 3D world's coordinate to the 3D camera's coordinate can be expressed by:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \underbrace{\begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix}}_{\mathbf{R}} \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + \underbrace{\begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}}_{\vec{t}}. \tag{1}$$

If f refers to the focal length of the camera, the geometric pinhole can be written as:

$$x = f\frac{x}{z} \quad \text{and} \quad y = f\frac{y}{z}. \tag{2}$$

Intrinsic camera parameters typically have an effective focal length f, a scale factor s, and an image center (u_0, v_0), which are also called a principal point. As usual, in the literature on computer vision, the origin of the image coordinate system is located at the upper left corner of the image. The unit of image coordinate is in pixels. For example, (x_{im}, y_{im}) is a pixel coordinate and (O_x, O_y) is the optical center. If the scale factors along the direction of x and y axis are s_x and s_y, respectively, then:

$$x = (x_{im} - O_x)s_x, \quad \text{and} \quad y = (y_{im} - O_y)s_y. \tag{3}$$

Based on Equations (2) and (3), we obtain:

$$\frac{f}{s_x}\frac{x}{z} = x_{im} - O_x \quad \text{and} \quad \frac{f}{s_y}\frac{y}{z} = y_{im} - O_y, \tag{4}$$

which implies:

$$x_{im} = \alpha_x \frac{x}{z} O_x \quad \text{and} \quad y_{im} = \alpha_y \frac{y}{z} O_y, \tag{5}$$

where α_x and α_y are considered as parameters for scaling at the direction of x and y.

In this method, a linear transformation equation can be made to map the world coordinate (x_w, y_w, z_w) to pixel coordinates (x, y) as:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \sim \underbrace{\underbrace{\begin{bmatrix} fs_x & fs_\theta & O_x \\ 0 & fs_y & O_y \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{K}} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\mathbf{M}_0} \begin{bmatrix} \mathbf{R} & \vec{t} \\ 0 & 1 \end{bmatrix}}_{\mathbf{M}} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \tag{6}$$

and therefore:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \sim \mathbf{M} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}. \tag{7}$$

Showing the elements of the composite matrix $\mathbf{M}$:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \sim \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}. \tag{8}$$

For a given constant α, we have:

$$\begin{bmatrix} \alpha x \\ \alpha y \\ \alpha \end{bmatrix} = \begin{bmatrix} m_{11}x_w + m_{12}y_w + m_{13}z_w + m_{14} \\ m_{21}x_w + m_{22}y_w + m_{23}z_w + m_{24} \\ m_{31}x_w + m_{32}y_w + m_{33}z_w + m_{34} \end{bmatrix}, \tag{9}$$

and therefore:

$$x = \frac{m_{11}x_w + m_{12}y_w + m_{13}z_w + m_{14}}{m_{31}x_w + m_{32}y_w + m_{33}z_w + m_{34}}, \text{ and} \tag{10}$$

$$y = \frac{m_{21}x_w + m_{22}y_w + m_{23}z_w + m_{24}}{m_{31}x_w + m_{32}y_w + m_{33}z_w + m_{34}}. \tag{11}$$

In the end, two equations with 12 unknown parameters $(m_{11}, m_{12}, \ldots, m_{34})$ are obtained. This is an over-determined system of linear equations. There are many ways to solve such system of linear equations, one of which is by using least-squares method [11]. In this paper, we will use singular value decomposition to solve such systems of linear equations.

B. Parameters Extraction using SVD

Extraction of parameters $\mathbf{M}$ cannot be solved directly, because the system of equations is over-determined. However, over-determined systems can be solved using SVD. In this method, the singular matrix $\mathbf{M}$ can be decomposed into:

$$\mathbf{M} = \mathbf{USV}^\top. \tag{12}$$

Since the vectors in $\mathbf{V}$ provide solutions related to the smallest eigenvalues, they provide real solutions. This way, parameters encoded in $\mathbf{M}$ are then extracted.

IV. Results and Analysis

A. Experimental Procedures

Two checkerboard patterns are pasted on the walls angled perpendicularly to one another as depicted in Fig. 3. Axes for the frame of the world are selected with the following steps: The point of origin is in the most bottom part, where two images meet at that angle. Axis Z moves upwards the point of origin. Axis X moves to the left wall of the point of origin. Axis Y moves to the right wall of the point of origin. The point of origin and three axes, namely, X, Y, and Z are shown in Fig. 3.

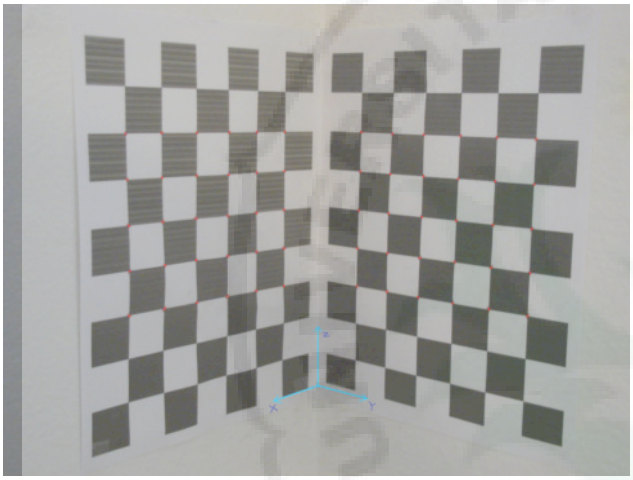

Fig. 3. Checkerboard pattern on the walls angled perpendicularly to one another and the coordinate axes for the frame of the world.

Points measured are shown with red dots on the image. A total of 60 points are selected, *i.e.*, 30 points in the left-hand checkerboard and 30 points in the right-hand checkerboard as shown in Fig. 3. Coordinates of the real world for the 60 points are measured using a ruler with a unit of centimeters.

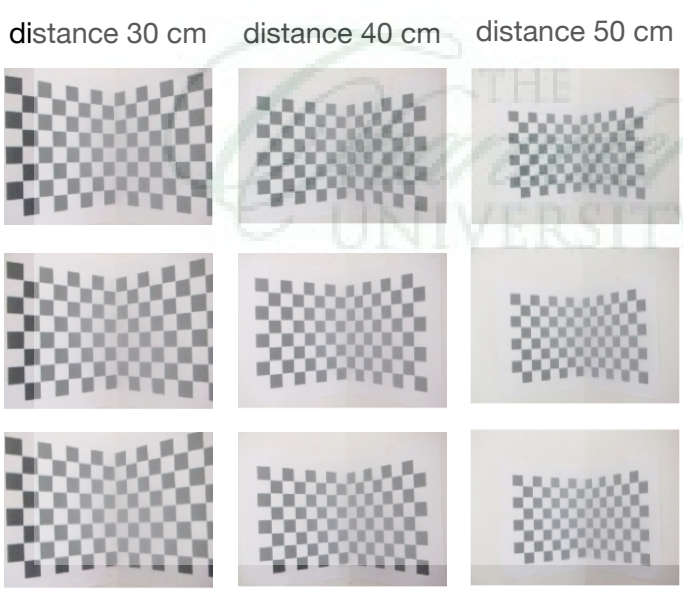

Fig. 4. Checkerboards captured from different distances.

Experiments are performed for three different distances from the camera to the object, *i.e.*, at a distance of 30 cm, 40 cm, and 50 cm as depicted in Fig. 4. Measurement for each distance is then repeated three times. This is to confirm the intrinsic parameters' values to be generated. Table I presents the resulting values for the intrinsic parameters generated from the experiments of camera calibration at different distances.

B. Results of the Calibration

Based on the results of the experiments, the following intrinsic parameters are obtained: the theta value is approximately $\frac{\pi}{2}$ (see Table I), meaning that the camera coordinate is not too skewed, and hence, the axes X and Y that are in the image frame are perpendicular to one another. The image resolution used in the experiments is 2592×1944, meaning that the center of the image is supposed to be at $(1296, 972)$. Nevertheless, experiments obtain $u_0 = 1701.2$ and $v_0 = 1592.3$ for the distance of 30_1, $u_0 = 794.1$ and $v_0 = 1093.5$ for the distance of 40_1 and so on. As measured in the experiments, the center of the image does comply with the principal point C_0 and differs $(405.2, 620.3)$ for the distance of 30_1.

TABLE I. The Resulting Intrinsic Parameters Obtained from Experiments of Camera Calibration using Checkerboards

Distance	Theta	u_0	v_0	Alpha (kf)	Beta (lf)
30_1	15.910 rad (90°)	1701.2	1592.3	2155.7	2062.1
30_2	15.789	794.1	1093.5	5057.3	4748.6
30_3	15.764	803.8	1202.7	5069.3	4764.6
40_1	15.787	1056.4	1071.4	5206.2	4921.4
40_2	16.343	976.0	1814.2	3406.9	3204.2
40_3	15.738	733.6	1108.2	5187.4	4885.9
50_1	15.762	1022.2	1012.7	5398.7	5108.4
50_2	15.746	833.7	1092.9	5440.6	5141.6
50_3	15.764	788.1	1083.4	5451.5	5159.6

Alpha and beta take measurement units of kf and lf, where k and l represent the number of pixels per centimeter, and f refers to the distance of the image frame physically from the equivalent pinhole or lens. Alpha and beta measurement units constitute the basis of the view point. It is revealed that the number of pixels per centimeter is 37.795275591. If $alpha = kf = 2155.7$ where $k = 37.795275591$, then $f = 57.03623$ is obtained and if $beta = lf = 2062.1$ where $l = 37.795275591$, then $f = 54.55973$ is obtained.

Based on the results of the experiment, the focal length obtained for each distance is given in Table II.

TABLE II. Focal Lengths Obtained in the Experiments for Various Distances

Distance	Left f	Right f
30 cm	134	126
40 cm	137	130
50 cm	143	136

The output of the extrinsic parameters for images at a

distance of 30_1 is as follows:

$$\mathbf{R} = \begin{bmatrix} -0.5077 & 0.8614 & -0.0130 \\ 0.1941 & 0.0997 & -0.9759 \\ -0.8394 & -0.4980 & -0.2179 \end{bmatrix}, \text{ and}$$

$$\vec{t} = [-4.0121 \quad -0.0267 \quad 28.9814].$$

Using the approach of least squares or SVD, rotation and translation matrices are obtained from the global coordinate system. Then, using the transformation matrices $\mathbf{R}$ and $\vec{t}$ obtained from each of those cameras (which are in their own local coordinate systems), the local coordinate systems can be translated into the same global coordinate system.

C. Application to Leaf Images

The values of the intrinsic camera parameters such as focal length (f) can be used to determine the depth of an image (z). If the disparity value (d) of a stereo image and a camera baseline (B) is known, the depth value of each point of the image can be determined by:

$$z = \frac{fB}{d}. \tag{13}$$

The basic concept of disparity is shown in Fig. 5, where P is a random point in three dimensions, which is considered as an object viewed from two cameras positioned in the same direction. Meanwhile, the two cameras may have different baselines. Furthermore, the object will appear in different positions in both stereo images.

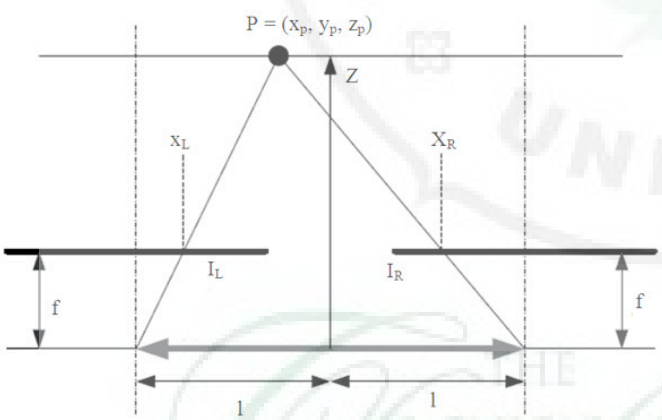

Fig. 5. A simple stereo system [12].

Disparity (d) is defined as the distance between the same object on the left and right stereo images, which can be expressed by:

$$d = x_L - x_R = f\left(\frac{x_p + 1}{z_p} - \frac{x_p - 1}{z_p}\right), \tag{14}$$

where, x_L and x_R are the coordinates of object x on the left and right stereo images [13]. From the disparity values of all points, we can construct a disparity image. An example of a disparity image can be seen in Fig. 6 [14].

Fig. 7 depicts the result of a 3D leaf-image reconstruction using disparity values obtained from correspondence calibration of stereo images and using baseline $B = 75$mm and focal length $f = 136$ [14].

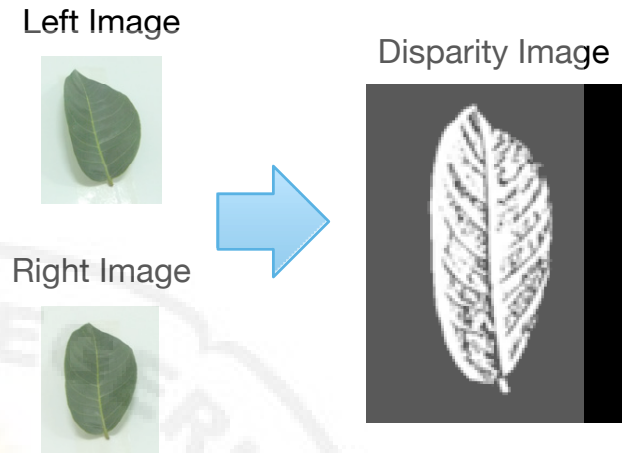

Fig. 6. The disparity image of a stereo leaf image.

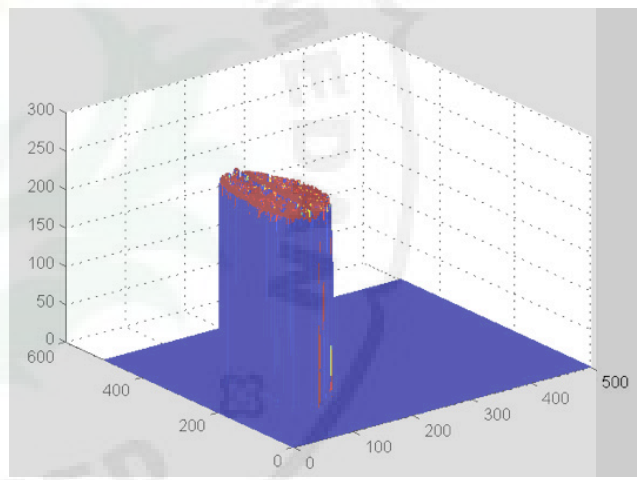

Fig. 7. Depth-map reconstruction based on disparity values of a leaf image with $f = 136$ and $B = 75$mm.

The results obtained show a rough surface, but the 3D leaf-surface results indicate that intrinsic parameters of camera calibration greatly contribute to building a disparity or depth map of a 3D leaf object. Fig. 7 in detail shows the depth of each pixel coordinate of the 3D leaf image. To obtain smoother depth map, it is necessary to perform smoothing on the resulting depth map or we can also use image rectification and segmentation [6]. Another solution can be done, namely, by refining calibration parameters through a non-linear optimization. The solution is obtained through minimizing an algebraic distance, which is not physically meaningful. This can be done through a maximum likelihood inference, which can be solved with the Levenberg-Marquardt algorithm [15].

D. Discussion

If the 3D point of the world model and the pixel coordinates of 2D projections are known, calibration matrices, rotation matrices, translation vectors, parameters for a scale at X and Y, and the optical center of the image (x, y) can be solved using a linear method with SVD, if the 3D point of the world model and pixel coordinates of 2D projections are known. By using SVD, the 3D point can be re-projected and by comparing

with the 2D point, the average pixel errors in the direction of X and Y can be revealed. For some experiments, different intrinsic values of focal length are generated. f values that can be used to find the depth are obtained from the mean of the obtained f values. Then, we have presented an experiment of depth map reconstruction using the intrinsic parameters of a camera. The results obtained show a rough surface compared with the results of this study [13]. Therefore, it is necessary to perform smoothing on the resulted depth map.

Although the results obtained show a rough surface, the 3D leaf-surface results indicate that intrinsic parameters (such as focal length) of camera calibration greatly contribute to building a disparity or depth map of a 3D leaf object. The depth maps generated in this study show nearly the same quality compared to the depth map of color imagery with region-based stereo image algorithms [16], even though the object's focus (*i.e.*, a leaf object) in this study tends to be small objects and, hence, to distinguish the coordinate points or object pixels on the image is more difficult. This suggests that the method proposed in this study may work well.

V. Conclusion

This paper has presented a simple approach to the reconstruction of the leaf image. Using SVD, calibration matrices, rotation matrices, translation vectors, parameters for a scale at X and Y, and the optical center of the image (x, y) can be obtained if the 3D point of the world model and pixel coordinates of 2D projections are known. SVD can also note the average pixel errors in the direction of X and Y by projecting a 3D point by comparing the return of 2D points. The 3D layout of the leaf image can be reconstructed into a 2D leaf image. Camera calibration is performed by observing a calibration object where the geometry of the 3D space is known with highly accurate precision. The intrinsic parameters of camera calibration greatly contribute in constructing disparity map or depth map of 3D leaf images. Furthermore, the results of the depth map reconstruction using intrinsic parameters of the camera show a rough surface making it necessary to perform essential post-processing steps, such as image rectification and segmentation, *i.e.*, smoothing the results of the depth map.

References

[1] H. J. Andersen, L. Reng, and K. Kirk, "Geometric plant properties by relaxed stereo vision using simulated annealing," *Computers and Electronics in Agriculture*, vol. 49, no. 2, pp. 219–232, 2005.

[2] S. Ericson and B. Åstrand, "Visual odometry system for agricultural field robots," in *Proceedings of the World Congress on Engineering and Computer Science*. International Association of Engineers, 2008, pp. 619–624.

[3] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, Nov 2000.

[4] M. Fiala and C. Shu, "Self-identifying patterns for plane-based camera calibration," *Machine Vision and Applications*, vol. 19, no. 4, pp. 209–216, 2008.

[5] J. Salvi, X. Armangué, and J. Batlle, "A comparative review of camera calibrating methods with accuracy evaluation," *Pattern Recognition*, vol. 35, no. 7, pp. 1617 – 1635, 2002.

[6] J. Weng, P. Cohen, and M. Herniou, "Camera calibration with distortion models and accuracy evaluation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 10, pp. 965–980, Oct 1992.

[7] L. Zhu, G. Ma, Y. Mu, and R. Shi, "Reconstruction 3D-models of old Beijing city by structured light scanning," in *Proceedings of the 22nd CIPA Symposium, 11-15 October 2009, Kyoto, Japan*, vol. 94, 2009.

[8] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International Journal of Computer Vision*, vol. 47, no. 1, pp. 7–42, 2002.

[9] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, "A comparison and evaluation of multi-view stereo reconstruction algorithms," in *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 1*, ser. CVPR '06. Washington, DC, USA: IEEE Computer Society, 2006, pp. 519–528.

[10] A. Chaudhury, A. Gupta, S. Manna, S. Mukherjee, and A. Chakrabarti, "Multiple view reconstruction of calibrated images using singular value decomposition," *arXiv preprint arXiv:1011.0596*, 2010.

[11] C. Gui and L. Tu, "A stereo camera calibration based on robotic vision," in *Cognitive Informatics Cognitive Computing (ICCI*CC), 2011 10th IEEE International Conference on*, Aug 2011, pp. 318–323.

[12] G. Wang and H. Ju, "A disparity map extraction algorithm for lunar rover BH2," in *Intelligent Computing and Intelligent Systems, 2009. IEEE International Conference on*, vol. 4, Nov 2009, pp. 385–389.

[13] P. Kamencay, R. Jarina, M. Breznan, M. Zachariasova, and P. Lukac, "Improved depth map estimation from stereo images based on hybrid method," *Radioengineering*, vol. 21, no. 1, pp. 70–78, 2012.

[14] H. Syahputra, A. Harjoko, R. Wardoyo, and R. Pulungan, "Plant recognition using stereo leaf image using gray-level co-occurrence matrix," *Journal of Computer Science*, vol. 10, no. 4, pp. 697–704, 2014.

[15] J. J. Moré, "The Levenberg-Marquardt algorithm: implementation and theory," in *Numerical analysis*. Springer, 1978, pp. 105–116.

[16] B. B. Alagoz, "Obtaining Depth Maps From Color Images By Region Based Stereo Matching Algorithms," Tech. Rep. arXiv:0812.1340, Dec 2008.