Computer Vision Techniques for Improving Structured Light Vision Systems

Yaan Zhang[1], Zhankun Luo[1], Jintao Hou[1], Lizhe Tan[1], Xinnian Guo[2]
[1]Department of Electrical and Computer Engineering, Purdue University Northwest, Indiana, USA
[2]Faculty of Electronic Information Engineering, Huaiyin Institute of Technology, Huaian, China
Email: zhan2722@pnw.edu;luo333@pnw.edu;hou100@pnw.edu;lizhetan@pnw.edu;yeamy1987@126.com

***Abstract*— In this paper, we propose computer vision techniques for 3D reconstruction and object height measurement using a single camera and multi-laser emitters which have an intersection on the projected image plane. Time-division and color division methods are first investigated. Although the color division method offers better accuracy for height measurement, it requires the laser emitters equipped with different color light, and the color division method is also sensitive to light exposure in the measurement environment. Next, a new multi-level random sample consensus (MLRANSAC) algorithm has been developed. The proposed MLRANSAC method not only offers high accuracy for height measurement but also eliminates the requirement for the laser emitters with different colors. Our experiment results have validated the effectiveness of the MLRANSAC algorithm.**

***Keywords*—Structured light vision; RANSAC.**

I. INTRODUCTION

Structured light has been widely applied in computer vision applications as a three-dimension reconstruction technique [1]-[3]. Comparing with the passive vision method, the structured light can adapt to less texture and low illuminate conditions, so it has higher reliability and accuracy over short distances. The structured light techniques also contribute a lot in industries such as archaeology, agriculture, architecture, autopilot and so on [4]-[7]. Laser emitters [1]-[3], [8]-[11] capable of emitting laser lines and projectors [10], [12]-[14] capable of posting encoded patterns are often used as light sources to utilize their geometric properties. A vision system with a light source and a camera placed at various angles provides sufficient information for a basic 3D reconstruction.

Since the image captured by a camera is the only source of information from the system and to reduce the difficulty of the feature extraction, the complexity of posting patterns and laser planes is limited. A common technique prefers parallel lines [4] rather than intersections in the structured light patterns to avoid the difficulties of image processing. For example, a laser-scanned system uses one single laser line and then reconstructs the model by moving the object along a certain orientation [8] or rotating the system along with a certain point [3], [7]. For a projector system, the binary gray-code pattern [10], [12] is widely applied, which is an accepted time-domain encode method. This method encodes each pixel as gray code by constantly posting patterns with black and white strips. Square dots are also applied as a post pattern in a projector system [13].

However, in the situations where the intersection emitting lines are used, the traditional method may fail to perform the 3D-reconstruction due to the ambiguity of labeling points around the intersection points [15]. In this paper, we aim to tackle the intersection line pattern captured in a structured light system equipped with multiple laser emitters. We will classify the unlabeled extracted laser points into different sets by utilizing the random sample consensus algorithm (RANSAC) and the results are compared with the time-division and color-division methods.

II. STRUCTURED LIGHT VISION SYSTEM WITH MULTI-LASER EMITTERS

Fig. 1 shows our measurement system. It consists of a high definition camera, a processing platform, and N laser transmitters with red, green, and blue color options. As shown in Fig. 1, the world coordinate and the camera coordinate are represented by $oxyz$ and $o_c x_c y_c z_c$, respectively. Note that the equation: $z = 0$ presents the horizontal plane in the world coordinate. It is assumed that the projection of the triangular plane on the plane of $z = 0$ exists and can be captured by the camera.

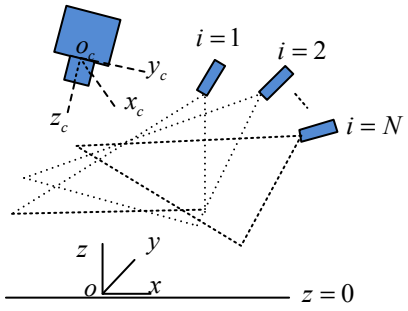

Fig. 1. Height measurement system.

A. Measurement Methods

The camera pinhole model is employed and its perspective projection is given below:

$$X_c = z_c A^{-1} \overline{I}_p \tag{1}$$

$$\overline{X}_c = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \overline{X} \tag{2}$$

For convenience, the following definitions are adopted:

978-1-7281-5317-9/20/$31.00 ©2020 IEEE

Authorized licensed use limited to: Purdue University. Downloaded on February 20,2021 at 08:19:32 UTC from IEEE Xplore. Restrictions apply.

$X = (x \quad y \quad z)^T$ - a point in the world coordinate

$X = (x \quad y \quad z \quad 1)^T$ - homogeneous form of X

$X_c = (x_c \quad y_c \quad z_c)^T$ - a point in the camera coordinate

$\bar{X}_c = (x_c \quad y_c \quad z_c \quad 1)^T$ - homogeneous form of X_c

$I_p = (u \quad v)^T$ - a point in the image pixel coordinates

$\bar{I}_p = (u \quad v \quad 1)^T$: homogeneous form of I_p

Note that s is a nonzero scale factor and A^{-1} intrinsic matrix A, which is defined below:

$$A = \begin{bmatrix} \alpha & c & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

where α and β are the projection of the effective focal length in x- and y- axes. u_0 and v_0 are the image intersection coordinates between the image plane and optical axis. c is the skewness in terms of two image axes. $[R \quad t] = [r_1 \quad r_2 \quad r_3 \quad t]$ designates the rotation and translation between the world coordinates and camera coordinates.

Let a characteristic point j on the triangular plane i formed by the ith laser emitter be:

$$\bar{X}_{cL}(i,j) = (x_{cL}(i,j) \quad y_{cL}(i,j) \quad z_{cL}(i,j) \quad 1)^T$$

and the ith plane is denoted by

$$\pi_i = (a_i \quad b_i \quad c_i \quad -1)^T .$$

Then it leads to the following:

$$X_{cL}(i,j)/z_c(i,j) = A^{-1}\bar{I}_p(i,j) \tag{4}$$

$$\pi_i \bar{X}_{cL}(i,j) = 0, \quad j = 1,2,\cdots,J \tag{5}$$

Given projected points $\bar{I}_p(i,j)$ and (4), we can obtain

$$X_{cL}(i,j)/z_{cL}(i,j) = \left[x_{cL}(i,j)/z_{cL}(i,j) \quad y_{cL}(i,j)/z_{cL}(i,j) \quad 1\right] \tag{6}$$

Given the calibrated triangular plane i, and (5), we finally can yield

$$z_{cL}(i,j) = 1/(a_i x_{cL}(i,j)/z_{cL}(i,) + b_i y_{cL}(i,j)/z_{cL}(i,j) + c_i) \tag{7}$$

After applying Equation (1), $X_c(i,j) = X_{cL}(i,j)$ can be obtained. Finally, we achieve

$$\bar{X}(i,j) = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix}^{-1} \bar{X}_c(i,j) \tag{8}$$

With the calibration information of extrinsic and intrinsic matrices, and the all the calibrated triangular planes, we can develop the time division, color division, and random sample consensus (RANSAC) method to conduct 3-D reconstruction.

B. Calibration

The camera is modeled using Zhang's method [17], where the augmented 2D and 3D points in the word coordinate are denoted by $\tilde{m} = (u \quad v \quad 1)^T$, $M = (x \quad y \quad z \quad 1)^T$, respectively. With a pinhole camera model, a relationship between a 3D point and its image project is given by

$$s\tilde{m} = A\,[R\ t]\ M = A\,[r_1\ r_2\ r_3\ t]\ M \tag{9}$$

Using $z = 0$ in the model plane (floor plane), (9) becomes

$$s\tilde{m} = A\,[r_1\ r_2\ t]\ \tilde{M} \tag{10}$$

where s is a scale factor and $\tilde{M} = (x \quad y \quad 1)^T$. The effective solutions for extrinsic R and intrinsic matrix A using (10) can be found in [17].

Each triangular plane $\pi_i = (a_i \quad b_i \quad c_i \quad -1)$ required in (5) for $i = 1,2,\cdots,N$ is calibrated independently. As shown in Fig. 1, let the model plane $z = 0$ in the word coordinate be $(0 \quad 0 \quad 1 \quad 0)^T$. We can translate the model plane to the camera coordinate as

$$\pi_0 = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix}^{-T} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} = [A \quad B \quad C \quad D]^T \tag{11}$$

As matter of fact, the projected J image points $X_c(i,j)$ in terms of the camera coordinates must be on the model plane of the checkerboard as well as in the model plane during calibration, that is,

$$\pi_0^T X_c(i,j) = 0 \ \text{ for } j = 1,2,\cdots,J \tag{12}$$

This is equivalent to

$$z_c(i,j)[Ax_c(i,j)/z_c(i,j) + By_c(i,j)/z_c(i,j) + C] = -D \quad \text{for } j = 1,2,\cdots,J \tag{13}$$

Note that $x_c(i,j)/z_c(i,j)$ and $y_c(i,j)/z_c(i,j)$ in (13) can be obtained using the projected pixel points, that is,

$$X_c(i,j)/z_c(i,j) = A^{-1}\bar{I}_p(i,j)) \ \text{ for } j = 1,2,\cdots,J \tag{14}$$

Then we can compute $z_c(i,j)$ in the camera coordinate via (13). Finally, we can achieve characteristic points as follows.

$$X_{cL}(i,j) = X_c(i,j) = z_c(i,j)A^{-1}\bar{I}_p(i,j) \tag{15}$$

Since the characteristic points are known on for $\pi_i = (a_i \quad b_i \quad c_i \quad -1)$, that is,

$$\pi_i \bar{X}_{cL}(i,j) = 0, \ j = 1,2,\cdots,J \tag{16}$$

and using M trials for different checkerboard orientation on the model plane at $z = 0$, it leads to the following conditions:

$$\pi_i \bar{X}_{cL}(i,j,k) = 0, \ j = 1,2,\cdots,J, \ k = 1,2,\cdots,M \tag{17}$$

Finally, the least squares method is applied to find the solution for triangular plane i, that is, $\pi_i = (a_i \quad b_i \quad c_i \quad -1)$.

III. IMPROVEMENT OF THE STRUCTURED LIGHT VISION SYSTEM

If a single camera shot is adopted and its image projection from the laser emitters has intersection, then the intersection points are difficult to be classified for measurement. The time-division, color division, and multi-level random sample consensus algorithm (MLRANSC) are proposed for applications.

A. Time Division Measurement

For the time division method, the laser emitters are operated sequentially at different times, that is, only one laser emitter can

Authorized licensed use limited to: Purdue University. Downloaded on February 20,2021 at 08:19:32 UTC from IEEE Xplore. Restrictions apply.

be turned on to hit on the surface of the measured object. This method is time-consuming.

After extracting the feature points corresponding to the projected image on the object and substituting the obtained feature points into the system model, the 3D information for these feature points can be calculated. Repeating this process for each laser emitter and composing the results from all the measurements, a 3-D reconstruction can be achieved.

B. Color Division Measurement

A color division method with a camera single shot is developed. By setting laser emitters with different colors, the laser planes are labeled physically. Therefore, we can separate the ambiguous intersection points by processing the corresponding color plane information in the captured image. The color threshold algorithm can be further applied to separate the color components from the laser emitter and the background color components.

Fig. 2 depicts the histograms of the measured in the RGB color space picture with turning on the red, green, and blue laser emitters simultaneously. As shown in Fig. 2, the projected pixels from the laser emitters are near the value of 255. The color threshold algorithm can be set as follows:

$$L_i = \begin{cases} 1 & Th_{low} \le p_i \le Th_{high}, p_j \le Th'_{high} \\ 0 & elsewhere \end{cases}, i \ne j \qquad (18)$$

where L_i is the binary matrix for a single laser with color i, p_i, p_j represents pixels of different color space dimension. Th_{low}, Th_{high} are thresholds while Th'_{high} eliminates the interference caused by other color dimensions. To optimize the color separating result, red, green and blue laser emitters are utilized to match the RGB color space in our work

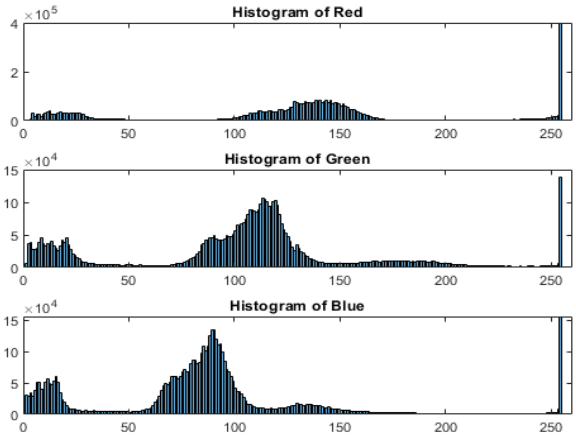

Fig. 2. Histogram of measured image in RGB color space.

This approach is effective, but it requires different color laser emitters. Besides, the color division method is sensitive to luminance, since with stronger luminance, parts of pixels which are not belong to lasers reach the higher value, even the maximum value. It affects the performance of the color threshold algorithm. To tackle this problem, the exposure time needs to be reduced before measurement.

C. Multi-Level Random Sample Consensus Algorithm (MLRANSAC)

In this method, we first add an additional step in the calibration stage. The points projected from the ith laser emitter on the horizontally placed checkerboard are extracted. The least squares algorithm can be applied to find its linear model, $\omega_i = (\lambda_1(i) \quad \lambda_2(i) \quad \lambda_0(i))^T$ corresponding to the plane $\omega_i(u(j) \quad v(j) \quad 1)^T = 0$. During the measurement stage at level 1 with a camera single shot, K points are randomly selected from the extracted point set. the normal vector is used initially. We compute K trial linear equations as following:

$$\overline{\lambda}_0(k) = -(\lambda_1(i)u_k + \lambda_2(i)v_k), \; k = 1, 2, \cdots, K \qquad (19)$$

Then we use a maximum allowable distance d to determine the number of points within allowable distance d. The selected line equation is the one consisting of the maximum number of points, that is,

$$\omega_i = (\lambda_1(i) \quad \lambda_2(i) \quad \lambda_0(i))^T, \; i = 1, 2, \cdots, N \qquad (20)$$

Fig. 3 depicts the scenario for case of $N = 2$.

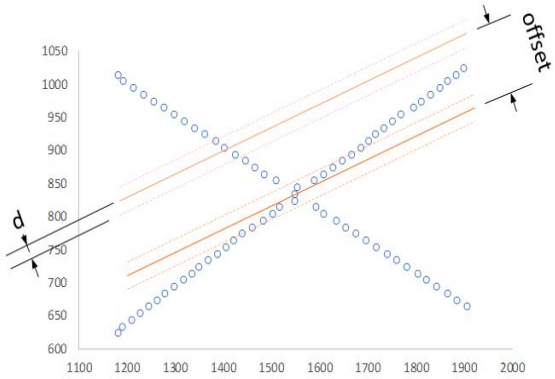

Fig. 3. First level of MLRANSAC.

Next, we start to use (20) as "the best straight line" representing the intersection of laser emitter i. We randomly choose two sample points from K point set to re-compute a straight line. With the obtained straight line, all extracted points are tested as follows. If the distance from a test point to the straight line is less than d, then the test point is considered on the intersection of the laser and added to the point set of the laser plane. When the cardinality of the laser plane point set corresponding to the straight line is greater than that of "the best straight line" and the angle between the straight line and the initial straight line in (20) within a threshold. Then "best straight line" will be replaced by the current one. Iterating K times this way, the final projected line equation is obtained. Repeating this process N, we achieve final N projected linear equations as well as N sets of the projected pixel points. Figs. 4-5 show the case of $N = 2$.

Finally, there are still some points which cannot be determined to which laser emitter they belong to (see Fig. 6). The distance factor and direction factor at level 3 are used. The distance factor is defined as a distance ratio (d_1 / d_2) of the undecided point from a test laser (red) intersection over the

Authorized licensed use limited to: Purdue University. Downloaded on February 20,2021 at 08:19:32 UTC from IEEE Xplore. Restrictions apply.

other laser (green) intersection. If the distance factor is less than a threshold, then the undecided point will be classified to this test laser point set. Fig. 6 describes for the case of $N = 2$.

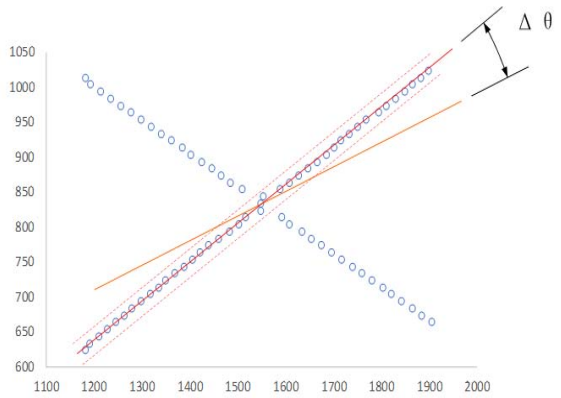

Fig. 4. Second level of MLRANSAC.

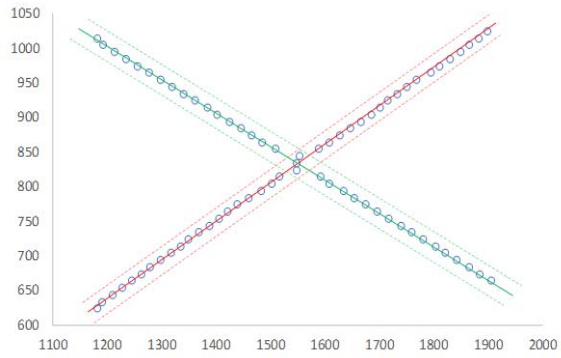

Fig. 5. The "best straight line".

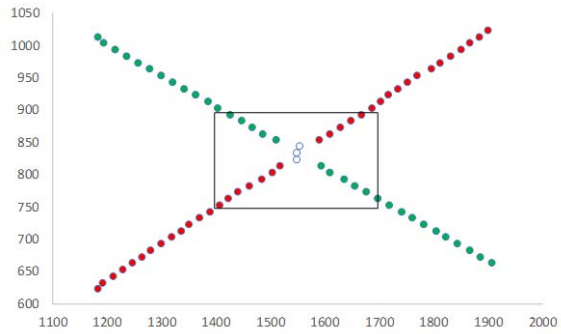

Fig. 6. Some points that cannot be decided to which laser emitters.

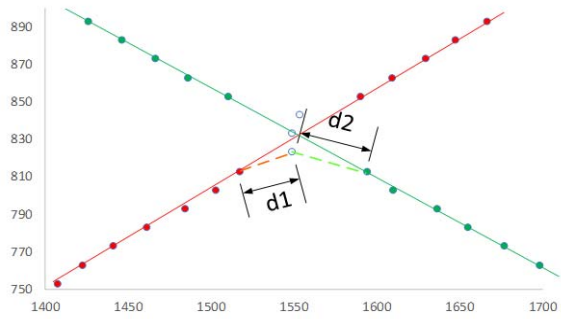

Fig. 7. The distance factor.

The direction factor is defined as a slope ratio of the slope between the undecided point and the nearest point in the text laser point set over the final direction of the projected line. The undecided point belongs to the test laser point set if the direction factor is closer to one among the other laser point sets, as shown in Fig. 8 the case of $N = 2$.

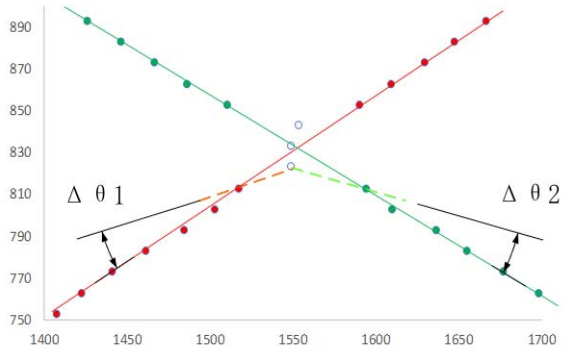

Fig. 8. The direction factor.

IV. Experiments and Processing

The proposed measurement system consists of a high definition camera (Basler acA2500-14gc GigE camera with ON Semiconductor MT9P031 CMOS sensor, 14 frames per second, 5MP resolution), a processing platform, checkerboard with 9*12 squares, laser transmitters with red, green, and blue color options. To validate our proposed algorithms, two laser emitters with color option are used. Fig. 9 depicts our system setup for measurement.

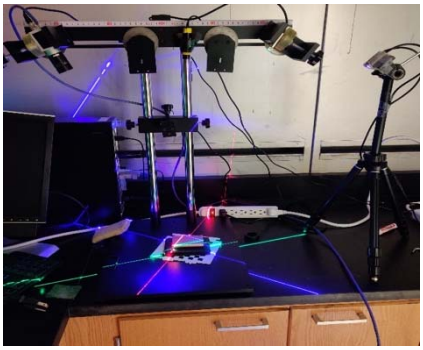

Fig. 9. Proposed structured-light measurement system

To eliminate random error caused by multiple experiments, we use red and blue lasers in a single experiment and use grayscale images as input of time division and MLRASAC algorithm. To calibrate the camera and lasers, fifteen poses of the checkerboard with lasers were captured as images separately.

A. System Calibration Results

With the Camera Calibration toolbox in MATLAB, intrinsic matrix A for the camera and extrinsic parameters $[R \quad t]$ are determined using Zhang's methodology [17]. The calibrated results are listed in (21).

Authorized licensed use limited to: Purdue University. Downloaded on February 20,2021 at 08:19:32 UTC from IEEE Xplore. Restrictions apply.

$$A=\begin{bmatrix} 7.603902\times10^{3} & 0 & 1.337963\times10^{3} \\ 0 & 7.585635\times10^{3} & 0.666364\times10^{3} \\ 0 & 0 & 1 \end{bmatrix}$$

$$R=\begin{bmatrix} -0.030498 & -0.999530 & -0.003148 \\ 0.939477 & -0.029741 & 0.341319 \\ -0.341252 & 0.007452 & 0.939942 \end{bmatrix} \quad (21)$$

$$t=\begin{bmatrix} 0.524791\times10^{2} \\ -0.149804\times10^{2} \\ 5.953130\times10^{2} \end{bmatrix}$$

Using the checkerboard images from the sequential laser emitter projections, we achieved two planes as below:

$$\pi_1=\begin{bmatrix} -1.895727\times10^{-3} \\ -3.302647\times10^{-3} \\ 1.924332\times10^{-3} \end{bmatrix} \quad \pi_2=\begin{bmatrix} 1.424982\times10^{-3} \\ -1.947035\times10^{-3} \\ 1.800564\times10^{-3} \end{bmatrix} \quad (22)$$

B. Height Measurement and Results

Figs. 10-12 display the objects for height measurements using our structured light vision system. During the measurement, we placed the object on the operating table to obtain six groups of object images, each group containing three pictures. When we used time division for height measurement, we selected images with single laser (Figs. 10-12 a and b). When using color division and MLRANSAC algorithm for height measurement, we used the images from two laser emitters as shown in Figs. 10-12 c.

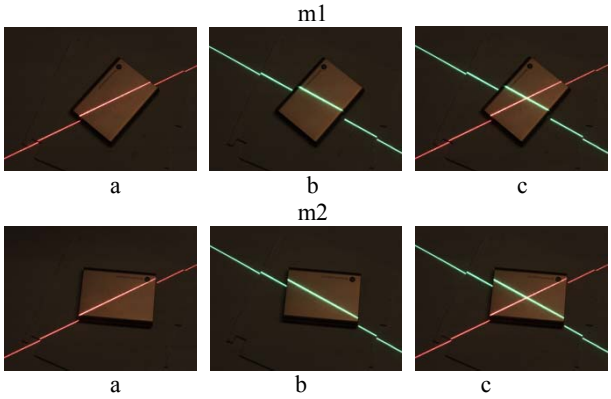

Fig. 10. Measurement for m1 and m2.

Tables 1-3 list the height measurement results. The MLRASAC algorithm has 6.70% relative error while the color division and time division offer 7.04% and 9.06%, respectively. The MLRANSAC algorithm offers the best results and also requires less process time when comparing with the time division and color division method and there is no specific color requirement for the laser emitter.

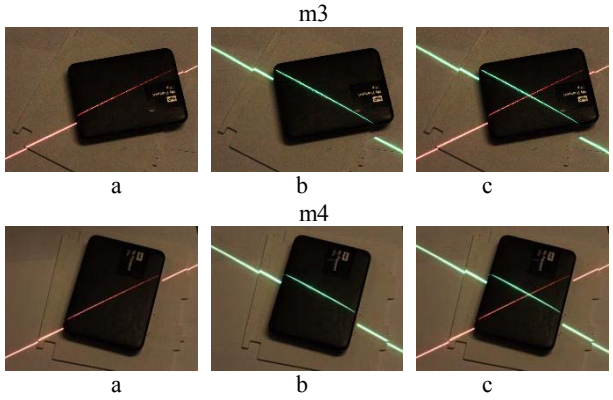

Fig. 11. Measurement for m3 and m4.

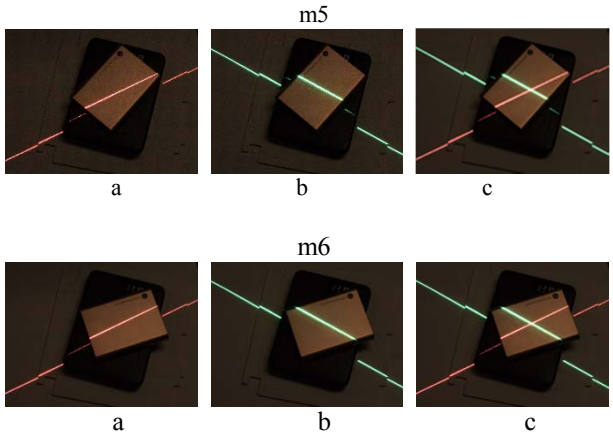

Fig. 12 Measurement for m5 and m6

Table 1 Time Division Method

Measurement Number	Actual height / mm	Measure height / mm	Absolute error / mm	Relative error / %
1	10.500	11.348	0.848	8.07%
2	10.500	11.421	0.921	8.77%
3	15.748	17.199	1.451	9.22%
4	15.748	17.414	1.666	10.58%
5	26.248	28.785	2.537	9.67%
6	26.248	28.358	2.110	8.04%

Table 2 Color Division Method

Measurement Number	Actual height / mm	Measure height / mm	Absolute error / mm	Relative error / %
1	10.500	11.507	1.007	9.59%
2	10.500	11.589	1.089	10.37%
3	15.748	16.648	0.899	5.71%
4	15.748	16.784	1.036	6.58%
5	26.248	27.504	1.256	4.79%
6	26.248	27.609	1.361	5.19%

Authorized licensed use limited to: Purdue University. Downloaded on February 20,2021 at 08:19:32 UTC from IEEE Xplore. Restrictions apply.

Table 3 Multi-level RANSAC algorithm

Measurement Number	Actual height / mm	Measure height / mm	Absolute error / mm	Relative error / %
1	10.500	11.324	0.824	7.85%
2	10.500	11.460	0.960	9.14%
3	15.748	16.839	1.091	6.93%
4	15.748	16.784	1.036	6.58%
5	26.248	27.655	1.407	5.26%
6	26.248	27.383	1.135	4.32%

C. 3-D Reconstruction Results

Figs. 13-15. display 3-D reconstructions uisng each of our proposed method. Fig.15. validates the effectiveness of MLRANSAC algorithm.

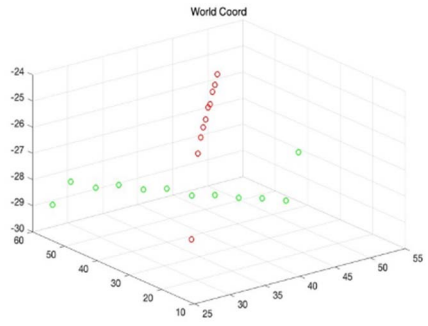

Fig. 13. 3-D reconstruction of m6 using time division.

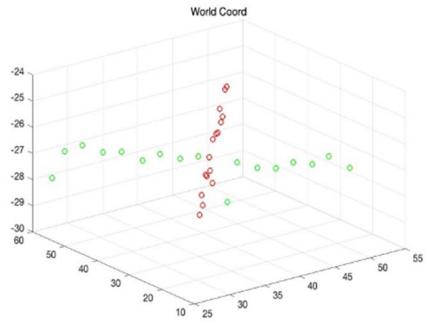

Fig. 14. 3-D reconstruction of m6 color division.

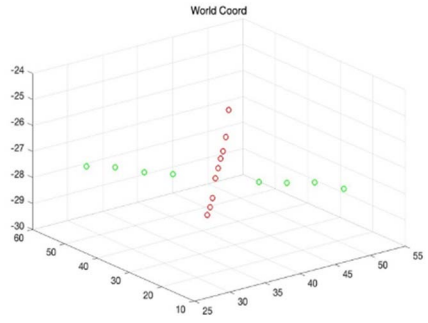

Fig. 15. 3-D reconstruction of m6 using the MLRANSAC algorithm

V. CONCLUSIONS

In this paper, we have comparatively investigated computer vision techniques for 3D reconstruction and object height measurement using a single camera and multi-laser emitters which have an intersection on the projected image plane. Proposed time division uses sequential operation on the laser emitters. Proposed color division method adopts the laser emitters with different colors. To improving the system, a new multi-level RANSAC (MLRANSAC) algorithm has been proposed. The algorithm not only offers high accuracy for height measurement but also eliminates the different color emitter requirement. Our experiment results have validated the effectiveness of the MLRANSAC algorithm.

REFERENCES

[1] G. Godin et al., "Laser range imaging in archaeology: issues and results," 2003 Conference on Computer Vision and Pattern Recognition Workshop, 2003.

[2] H. Ha, T.-H. Oh, and I. S. Kweon, "A multi-view structured-light System for highly accurate 3D modeling," 2015 International Conference on 3D Vision, 2015.

[3] C. Ho, "Machine vision based 3D scanning system," in 2009 9th International Conference on Electronic Measurement & Instruments, 2009, pp. 4-445-4-449: IEEE.

[4] J. Li, G. Liu, and Y. Liu, "A dynamic volume measurement system with structured light vision," in 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC), 2016, pp. 251-255: IEEE.

[5] M. Garrido, M. Perez-Ruiz, C. Valero, C. Gliever, B. Hanson, and D. Slaughter, "Active optical sensors for tree stem detection and classification in nurseries," Sensors, vol. 14, no. 6, pp. 10783–10803, 2014.

[6] D. Li, H. Zhang, Z. Song, D. Man, and M. W. Jones, "An automatic laser scanning system for accurate 3d reconstruction of indoor scenes," in 2017 IEEE International Conference on Information and Automation (ICIA), 2017, pp. 826-831: IEEE.

[7] J. Deng, B. Chen, X. Cao, B. Yao, Z. Zhao, and J. Yu, "3D reconstruction of rotating objects based on line structured-light scanning," in 2018 International Conference on Sensing, Diagnostics, Prognostics, and Control (SDPC), 2018, pp. 244-247: IEEE.

[8] J. Fan, F. Jing, L. Yang, L. Teng, and M. Tan, "A precise initial weld point guiding method of micro-gap weld based on structured light vision sensor," IEEE Sensors Journal, vol. 19, no. 1, pp. 322-331, 2019.

[9] S. R. Fanello et al., "HyperDepth: learning depth from structured light without matching," presented at the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[10] C. Holenstein, R. Zlot, and M. Bosse, "Watertight surface reconstruction of caves from 3D laser data," 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2011.

[11] X. Chen et al., "A structured-light-based panoramic depth camera," 2018 IEEE International Conference on Real-time Computing and Robotics (RCAR), 2018, pp. 102-107: IEEE.

[12] D. Scharstein and R. Szeliski. "High-accuracy stereodepth maps using structured light," In CVPR, 2003.

[13] M. Vo, S. G. Narasimhan, and Y. Sheikh, "Texture illumination separation for single-shot structured light reconstruction," IEEE Trans Pattern Anal Mach Intell, vol. 38, no. 2, pp. 390-404, Feb 2016.

[14] H. Ha, T.-H. Oh, and I. S. Kweon, "A multi-view structured-light System for highly accurate 3D modeling," presented at the 2015 International Conference on 3D Vision, 2015.

[15] L. Zhang, J. Lin, J. Sun, G. Yin, C. Ma, and L. Nie, "A robust stripe segmentation method for 3D measurement of structured light," 2015 IEEE International Conference on Mechatronics and Automation (ICMA), 2015, pp. 2431-2436: IEEE.

[16] Bouguet, J.Y. Camera Calibration Toolbox for Matlab; California Institute of Technology: Pasadena, CA, USA, 2013.

[17] Z. Zhang, "Flexible camera calibration by viewing a plane from unkown orientations", Proceedings of the Seventh IEEE International Conference on Computer Vision, September 1999.

Authorized licensed use limited to: Purdue University. Downloaded on February 20,2021 at 08:19:32 UTC from IEEE Xplore. Restrictions apply.