Stereo Camera – based 3D Object Reconstruction Utilizing Semi-Global Matching Algorithm

M.S. Hendriyawan Achmad[1,*], Widya Setia Findari[1], Nurnajmin Qasrina Ann[2], Dwi Pebrianti[2], Mohd Razali Daud[2]

[1]*Faculty of Science and Technology, Universitas Teknologi Yogyakarta (UTY), Yogyakarta, Indonesia*
[2]*Faculty of Electrical & Electronics Engineering, Universiti Malaysia Pahang (UMP), Pekan, Malaysia*
**contact: hendriyawanachmad@uty.ac.id*

***Abstract*—a 3D reconstruction using stereo cameras still becomes an issue among researchers specialized in computer vision. The corresponding pixel between two images from stereo camera needs to be estimated accurately. One of the widely used methods is Semi-Global Matching (SGM), which uses mutual information (MI) in the form of entropy between two pixels to determine the level of similarity based on the smallest energy (lower cost). The reconstruction result shows the percentage of registered pointcloud is equal to 62.11% where the observation distance ranges are between 1 to 4 meters. In this research, a nearest-neighbor filter is utilized to improve the pointcloud quality where the variations of the neighbor's number are 4 to 128 pixels. The results show that this technique can eliminate the outliers up to 4.9% with the standard deviation of nearest-neighbor distances means equals to 1.0.**

Keywords—stereo camera; 3D reconstruction; semi-global matching; pointcloud;

I. Introduction

In general, computer vision only sees a 2D object based on image color and intensity due to the reflection of the light sources illuminating the object. 3D object reconstruction based on stereo camera requires the disparity value of the corresponding pixels. The 2D computer vision techniques from the single camera cannot calculate the difference value. Recently, the implementation of 3D reconstruction technology is already widely needed at various areas, including geological mapping [1], [2], the inspection of products in the industry [3], and robotics [4], [5]. One of the techniques that has been known and often used for projecting a 3D object is a laser scanner (LIDAR), as has been done by [6], [7] which has a high degree of accuracy, but still costly to implement. A stereo vision technique to obtain 3D images using the concept of disparity is the method which low-cost among others.

3D reconstruction of the object using a computer stereo vision is about how the computer perceives the object of observation as seen from two different sides just like the human eye [8]. A stereo vision system uses two cameras, as shown in Fig. 1. Ideally, the two cameras are separated by a short distance and mounted almost parallel to each other. The stereo vision algorithms utilize the disparity of the two corresponding pixels of stereo images to calculate the depth value. Parameter calibration of the stereo cameras can be used to determine the 3D spatial relationship of corresponding pixels, resulting in a disparity map and provide more detailed data to identify objects, detect defective products, guide the robot movement toward a goal, and so forth. In this paper, we introduce a 3D object reconstruction method using stereo cameras utilize Semi-Global Matching algorithms (SGM) for projecting two stereo images into a single 3D image accurately. This technique is expected to become one of the methods for 3D object reconstruction which is cheap and reliable.

This paper is composed of few sections. Section 2 describes the stereo camera calibration methods to get the intrinsic and extrinsic parameters and the working principle of SGM algorithm to map the disparity in the framework of the 3D reconstruction. Section 3 describes the results of the calibration to obtain the low rate of re-projection error. Moreover, it is also presented the results of a 3D object reconstruction based on calibration parameters and the implementation of SGM algorithm. Section 4 describes the conclusions based on the results of 3D reconstruction experiments using a stereo camera utilizes the SGM algorithm.

II. Methods

A. Research Flow

Research on 3D reconstruction using stereo camera needs some critical stages. The first stage is the calibration of the stereo camera. The second stage is the acquisition and pre-processing of the 2D stereo images. The third stage is to build a map of disparities using SGM algorithm. The fourth stage is a 3D reconstruction using disparity map and calibration parameters. The last step is a refinement on 3D images to eliminate the outlier points using the nearest neighbor filter as it is done by [9], [10]. Fig. 2 gives an overview of the research flow from the initial stage to the final stage.

Fig. 1. The stereo camera

978-1-5090-4357-6/16/$31.00 ©2016 IEEE

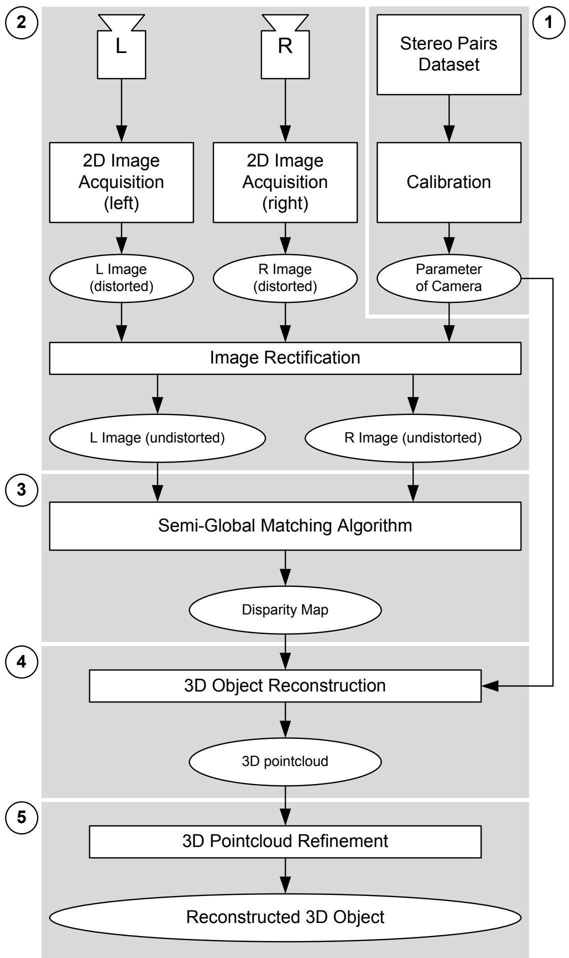

Fig. 2. The study flow on 3D object reconstruction using stereo camera

The calibration phase aims to obtain the information about the parameters of the camera and conducted in the first stage because the next step requires the camera parameter information as an input. The camera parameters are used to eliminate the radial and tangential distortions of 2D image acquisition which generated in the second stage (pre-processing). The radial distortion is caused by the deflection of rays on the camera lens, while the tangential distortion is caused by the lenses and the camera sensor which are not perpendicular. Ideally, there are no radial and tangential distortion types for pinhole cameras, because it does not use a lens. This study uses the camera which has a lens so that we consider the distortion coefficients in the calibration parameters to correct the two types of distortion. The third stage is the most critical step which determines the quality of the reconstruction. At this juncture, the two stereo images that contain no distortion are matched using the algorithm of Semi-Global Matching (SGM) at the pixel level [11] to generate the disparity map as a raw material for the 3D reconstruction stage. The fourth step is 3D reconstruction that requires the information of camera parameters to build a 3D image pointcloud in world coordinates (XYZ) based on the disparity map. The fifth stage is to improve the quality of the pointcloud using a nearest neighbor filter. The refinement result shows the recognizable reconstructed object.

TABLE I
SPECIFICATIONS OF THE CAMERA AND LENS

Type	Specifications	
Camera	Resolution	1600 x 1200
	Frame Rate	50 fps
	Pixels	2.0 MP
	Sensor	CMOS
	Chrome	Colour
Lens	Focal Length	15mm – 50mm
	Optical Format	1/3”

B. Camera Prototype

The camera also determines the quality of 3D reconstruction result. Two cameras are configured as a stereo camera and must be identical, which have the same hardware specifications, such as resolution, the number of pixels, and the type of sensor. Similarly, the use of lens must have the same specs including a focal length and optical formats. These specs are to synchronize the perception of two cameras to the object so that the result of 3D reconstruction is getting the best performance. Fig. 1 shows two identical cameras which are configured to be a stereo camera. Table I describes the specifications of the camera and lens.

C. Camera Calibration

Camera calibration is also known as camera resectioning. That is the process of getting the parameters of the sensor and the camera lens to capture the object images. These parameters are used to correct camera distortion (radial and tangential). In a stereo camera system, the parameters are used to project the object of observation in 3D coordinates, or specify the location of the camera in 3D coordinates about the images position in the real coordinate.

The camera parameters which obtained from calibration process consist of the intrinsic, extrinsic and distortion coefficient. In this study, camera calibration process uses a checkerboard as a calibration pattern, which has constructed from of 9x7 squares with an area of 52 mm^2 for each square as shown in Fig. 3.

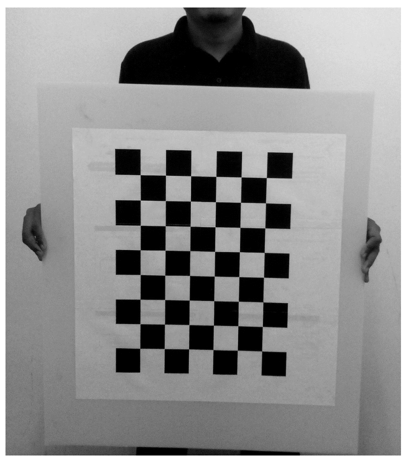

Fig. 3. The checkerboard for camera calibration

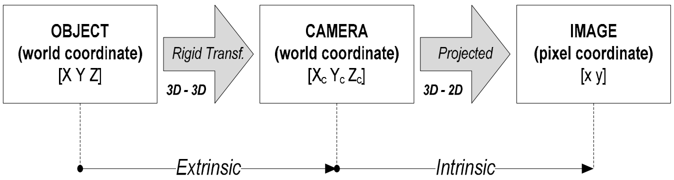

Fig. 4. The real object transformation into pixel

Apart from being used to obtain the camera parameters, the checkerboard is applied to determine the value of the re-projection error between the projection points to the real coordinate points. The value of the re-projection error indicates the quality of the calibration results. This experiment uses the 12 combinations of checkerboard image pairs as stereo camera calibration input with the variation of distance (translation) and orientation (rotation).

Mathematically, the camera parameters are written in the form of a matrix P which has two elements, extrinsic and intrinsic. While the matrix of distortion coefficients has its own, because the distortion coefficients are only used on the camera that has a lens. Since this study uses the camera with a lens, image correction process uses the camera distortion coefficients. The matrix model (P) of the camera is defined by (1):

$$P = K\begin{bmatrix} R & t \end{bmatrix} \tag{1}$$

R and t are the extrinsic parameters, which represent the camera rotation vector and the camera translation vector, respectively. Additionally, K is the intrinsic parameter. Fig. 4 illustrates the benefit of camera parameters to transform the real object in 3D coordinate about the camera coordinate in 3D coordinate as well, into a 2D image projection in the form of the disparity map.

Object transformation in the world coordinates into pixel coordinates as shown in Fig. 4 is defined by using (2) below:

$$s\begin{bmatrix} x & y & 1 \end{bmatrix}^T = K\begin{bmatrix} R & t \end{bmatrix}\begin{bmatrix} X & Y & Z & 1 \end{bmatrix}^T \tag{2}$$

[x y] is the pixel coordinate of the 2D image by a scaling factor s. It is obtained by multiplying the camera parameter with the world coordinates of the object [12].

1) The intrinsic parameter is the parameter of the camera which provides information of the sensor model including the value of focal length, optical point, and skew coefficient. Intrinsic parameter matrix (K) defined by (3):

$$K = \begin{bmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

(f_x , f_y) are the scale factors of the focal length of the camera, s is a skew coefficient of the image captured by the sensor, while (c_x , c_y) is the coordinate of the sensor midpoint (optical point/principal point). All values are in pixels unit.

2) The extrinsic parameter is the parameter of the camera which provides information about the rigid transformation of 2D projection images to the 3D coordinates (world). The rigid transformation consists of rotation and translation. Rotation function is defined by the matrix vector R (R_x, R_y, R_z) and a translation function is specified by the matrix vector t (t_x, t_y, t_o). The reference point of transformation is on the Z axis of the optical point camera.

3) Distortion Coefficient is the parameter of the camera which provides information about the distortion coefficients of the lens. There are two types of distortion coefficients, such as radial distortion and tangential distortion [13]. Radial distortion coefficient (k) indicates the intensity of the 2D image distortion due to the bending of light rays by the lens; the negative value indicates a pincushion distortion and the positive value indicates a barrel distortion. The coefficient of tangential distortion (p) shows the coefficient of alignment between the sensor and the camera lens. The coefficients are used to correct distortion in the 2D image so that the two types of distortion (radial and tangential) can be eliminated. This stage is part of the pre-processing stage as depicted in Fig. 2. To obtain the corrected pixel coordinates of the radial distortion (x_k , y_k) by using the equations denoted by (4)(5):

$$x_k = x\left(1 + k_1 r^2 + k_2 r^4 + k_3 r^6\right) \tag{4}$$

$$y_k = y\left(1 + k_1 r^2 + k_2 r^4 + k_3 r^6\right) \tag{5}$$

(x, y) is representing the original pixel coordinate of the input image which still affected by the distortion. (k_1, k_2, k_3) are the coefficients of the radial distortion, and $r^2 = x^2 + y^2$. Similarly, to obtain the corrected pixel coordinates of the radial distortion (x_p,y_p) using the equation defined by (6)(7):

$$x_p = x + \left[2p_1xy + p_2\left(r^2 + 2x^2\right)\right] \tag{6}$$

$$y_p = y + \left[p_1\left(r^2 + 2y^2\right) + 2p_2xy\right] \tag{7}$$

(p_1, p_2) are tangential distortion coefficients, and $r^2 = x^2 + y^2$. Thus the distortion coefficients are required to correct the 2D image from the radial and tangential distortion. Where the distortion coefficient are denoted by k_1, k_2, p_1, p_2, k_3.

D. Semi-Global Matching (SGM) Algorithm

This algorithm was firstly introduced by [11] to determines the similarity between the two images in the operation of pixels. In general, the corresponding pixels are based on the difference in intensity which sensitive to the lighting, the reflection, the recording process, and so forth. Semi-Global Matching (SGM) algorithm uses mutual information (MI) of the two images. The concept of mutual information has been used to match the stereo images based on correlation by [14] and Graph Cuts by [15]. The results show that the matching process is not sensitive to the differences in the lighting and reflections as occurs in the conventional matching method.

MI which used as the basis for the matching process is defined by the value of the entropy (H) of the two stereo images, as defined by (8):

$$MI_{I_1,I_2} = H_{I_1} + H_{I_2} - H_{I_1,I_2} \tag{8}$$

H_{I_1} and H_{I_2} are the entropy value of the pixels intensities (i) from the left image and the right image. While H_{I_1,I_2} is the joined entropy value of the two corresponding pixels. Entropy values are calculated from the distribution probability (P) of the associated image pixels, as defined by (9) below:

$$H_I = -\int_0^1 P_I(i)\log P_I(i)\,di \tag{9}$$

While the joint entropy value is calculated based on the method proposed by [15], as defined by (10) below:

$$H_{I_1,I_2} = \sum_p \left(-\frac{1}{n}\log\left(P_{I_1,I_2}(i_{1p}, i_{2p}) \otimes g(i_{1p}, i_{2p}) \right) \otimes g(i_{1p}, i_{2p}) \right) \tag{10}$$

n and $P_{I_1,I_2}(i_{1p}, i_{2p})$ define the number and the distribution probability of the pixel p which is matched in the epipolar line respectively. $\otimes g(i_1, i_2)$ is the 2D Gaussian convolution of the two pixels which are mutually matched.

Semi-Global Matching algorithm runs on every horizon epipolar of the two stereo images which have been corrected (rectified). Each step of matching calculates the value of matching cost based on the MI as defined in equation (2). Then the lowest value of matching price is chosen to determine the location of the corresponding pixels. Ideally, the SGM algorithm search for the corresponding pixels in each epipolar line of the stereo image, but it requires a significant computational time. Therefore, the matching can be done in some points of pixels coordinate, and the rest can be predicted in a linear way to get the disparity map of the stereo image. This matching method reduces a computational time, but it requires the input images which have low distortion values.

E. Disparity to 3D Image Transformation

The disparity image represents the distance of the two corresponding pixels of the left image and a right image. The disparity value is converted into depth value (Z axis of the camera) to get the 3D image pointcloud. These process needs the information in the form of camera parameters to change the disparity value into the depth value, as shown in Fig.2 stage-4. Z values are sought for each pixel (x,y) on the disparity image. The pointcloud (X, Y, Z) are determined based on the (11-13):

$$Z = f\frac{B}{D} \tag{11}$$

$$Y = (R - CR)\frac{Z}{f} \tag{12}$$

$$X = (C - CC)\frac{Z}{f} \tag{13}$$

f is the focal length, B is the baseline, D is the disparity, R is the row of the pixel, C is a column of the pixel, CR is the center row of the disparity image, and CC is the center column of disparity image. The accuracy of the estimated value Z can be measured based on the re-projection error value (e_p) which is obtained during the calibration process. Based on (11), assuming a depth value Z is the already known (ground truth), the value of D can be calculated. Thus, the accuracy of the estimated Z (ΔZ) can be calculated based on the equation defined by (14).

$$\Delta Z = \left| \left(f\frac{B}{D} \right) - \left(f\frac{B}{D - e_p} \right) \right| \tag{14}$$

Equation (14) can be used to evaluate the calibration parameters on the results of 3D reconstruction. If ΔZ provides value which cannot be accepted, the re-calibration phase should be performed.

III. Results and Discussions

A. The Calibration Results

The results of stereo camera calibration determine the quality of the 3D reconstruction. The value of re-projection error is used as the standard of quality of the calibration. For calibration purpose, a checkerboard is set in 12 different distances and orientations. Table II describes the intrinsic parameters of the camera obtained from the calibration process.

TABLE II
THE INTRINSIC PARAMETERS OF THE STEREO CAMERA

Camera	Parameter	Axis	Value	Std. Error
Left	Focal Length (pixel)	f_x	914.449	± 3.8374
		f_y	916.042	± 3.8101
	Optical Centre (pixel)	c_x	272.356	± 4.3413
		c_y	221.650	± 4.8599
	Skew (pixel)		1.063	± 0.5646
Right	Focal Length (pixel)	f_x	934.661	± 3.9254
		f_y	936.181	± 3.9108
	Optical Centre (pixel)	c_x	270.982	± 4.5488
		c_y	219.509	± 4.7821
	Skew (pixel)		0.859	± 0.5612

TABLE III
THE DISTORTION COEFFICIENTS OF THE STEREO CAMERA

Camera	Distortion	Coefficient	Value
Left	Radial	k_1	-0.3542
		k_2	0.065
		k_3	0.58
	Tangential	p_1	-0.0008
		p_2	-0.003
Right	Radial	k_1	-0.3472
		k_2	-0.101
		k_3	1.42
	Tangential	p_1	-0.0011
		p_2	-0.002

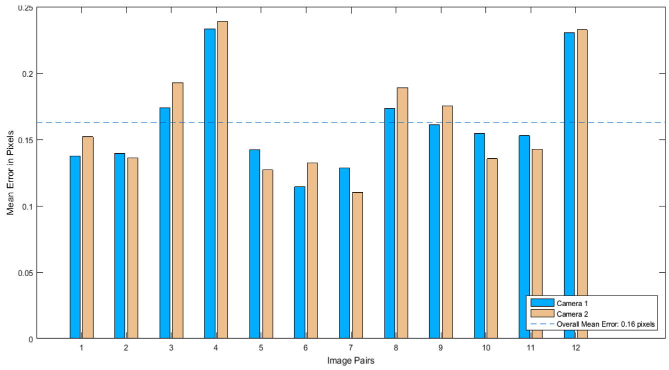

Fig. 5. The re-projection error rates of 12 checkerboard positions produce a mean error equals to 0.16 pixels.

Fig. 6. Distorted images correction. (left) The image before correction. (right) The image after correction.

Based on the values shown in Table II and refer to (3), the fundamental matrix (K) for each of the camera can be calculated. The stereo camera can project the 12 positions of the objects (checkerboard) in world coordinates about the coordinates of the camera based on translation and rotation matrix which have obtained. The radial and tangential distortion coefficients obtained from calibration stage are used to correct stereo images using an equation defined by (4-7) as shown in Table III.

The accuracy of the re-projection is defined by the mean value of re-projection errors. Fig. 5 shows the re-projection errors rates of 12 object positions during calibration. The result shows the average amount of re-projection errors is 0.16 pixels. With re-projection error value < 0.2 pixels, it can be concluded that the calibration results are good enough and the calibration parameters can be used for the next stage.

The results of stereo image correction based on the distortion coefficients shown in Table III can be evaluated from the corrected image as shown in Fig. 6.

B. Disparity Map

The disparity map is a 2D image which each pixel represents the difference of position between two pixels that match in the stereo image. The stereo image used to create the disparity map should be corrected first of all distortion (radial and tangential) before the matching process using SGM algorithm. The stereo image which has been corrected and rectified, make it eligible of the two images to be matched on each epipolar line. So that, the disparity value of the two corresponding pixels can be calculated in the right way.

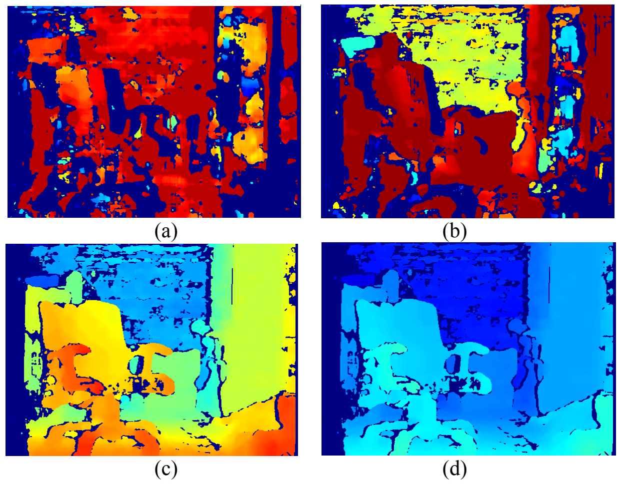

Fig. 7. The disparity maps for some ranges. (a) Disparity range = 0-16 pixels. (b) Disparity range = 0-32 pixels. (c) Disparity range = 0-64 pixels. (d) Disparity range = 0-128 pixels.

Fig. 7 shows the experimental results of generating the disparity maps within the interpretation ranges to get the best results approaching ground truth value. Fig. 7(c) shows the best disparity image than others, which uses the disparity range 0 to 64 pixels. The blue color indicates the region where the disparity values are not detected because corresponding pixels are not found in the area. The small value in disparities produces a high value of Z, even infinite (zero gaps). In contrast, large values in disparities provide a little value of Z, which marked with colors closer to the red. The location of the chair is closer to the camera than the bookshelf. Part of the seat which close to the camera tends to be a red, compared with the bookshelf tends to away from red.

Based on the result, this study concludes that the range of disparity value is calculated based on the maximum distance of the object which is observed, such as focal length, baseline, and the pixel size of the sensor as defined by (11).

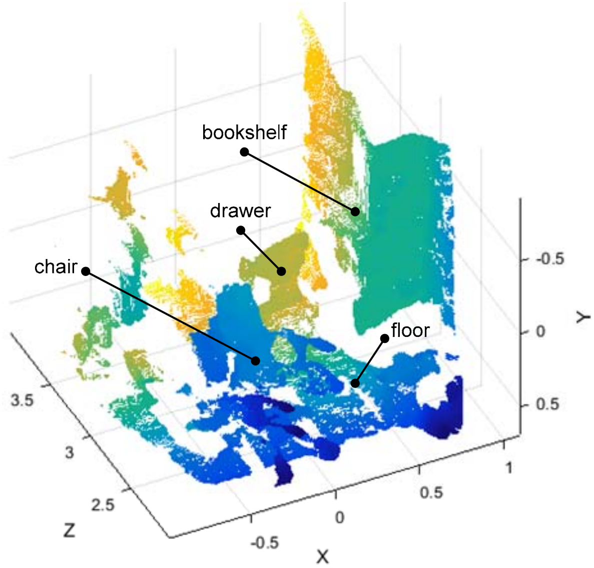

Fig. 8. The 3D image pointcloud after reconstruction.

TABLE IV
THE NUMBER OF NEIGHBORS VS THE PERCENTAGE OF INLIER AND OUTLIER

Pointcloud	Neighbour (pixels)	Inlier (pixels)	Outlier (pixels)	Inlier (%)	Outlier (%)
Original	0	192340	0	100	0
NN-004px	4	192175	165	99.9	0.1
NN-008px	8	191950	390	99.8	0.2
NN-016px	16	191078	1262	99.3	0.7
NN-032px	32	188937	3403	98.2	1.8
NN-064px	64	185946	6394	96.7	3.3
NN-128px	128	182972	9368	95.1	4.9

C. The Results of 3D Reconstruction

Once a good disparity map was obtained, the next is to project it into the 3D image pointcloud, which transforms the disparity map in pixels into pointcloud in meters by using (11-13) as shown in Fig. 8. Ideally, the results of 3D reconstruction produce a pointcloud as many as 309700 points, according to the total pixels of the stereo image which have been corrected by 475 x 652 pixels. However, not all disparities come up with an acceptable value, for example, the value is too large (-infinite or +infinite). So those points have to be removed. Fig. 8 shows the pointcloud which has been selected based on a range of distance (Z) between 1-4 meters. Selection in pointcloud produces registered points with the number of 192000 pixels and unregistered points with the number of 117360 pixels. Thus, the percentage of the registered points of 62.11% and the remaining of 37.89% is the unregistered points. The percentage of the registered points can still go down because pointcloud has not been cleaned from the outliers yet. Table IV shows the comparative results of the refinement stage against the pointcloud using a nearest-neighbor filter with variations in the number of neighbor between 4-128 pixels and the standard deviation value of the mean of nearest-neighbor distance = 1.0.

The experimental results of the pointcloud refinement utilizing nearest-neighbor filters by varying the amount of the neighbors, there is a rise in the percentage of outliers. However, the increase in the proportion of outliers is not proportional to the increase in the number of neighbors used to reduce outlier. Significant additions to the number of neighbors do not make a significant reduction in the percentage of outliers.

IV. CONCLUSIONS

The quality of the 3D reconstruction using stereo camera is very sensitive to the input parameters of the camera. A small error value of re-projection gives better reconstruction quality. In this study, we get the value of the re-projection error is 0.16 pixels, which is still below 0.2 pixels taken as a rule of thumb on the calibration of the stereo camera. 3D object reconstruction using a stereo camera which performs A Semi-Global Matching (SGM) algorithm is one of the methods to find the corresponding pixels, which is using the mutual information (MI) in the form of entropy between pixels to determine the level of similarity based on the smallest energy (lower cost). The reconstruction result shows the percentage of registered pointcloud is equals to 62.11% within the observation distance ranges between 1 to 4 meters. The pointcloud refinement method applies the nearest-neighbor filter; with the variations in the number of neighbors between 4-128 pixels. The result shows that it can eliminate the outliers up to 4.9%.

ACKNOWLEDGMENT

This research was jointly supported by the Junior Lecturer Research Grant (013/HB-LIT/III/2016) from the Ministry of Research Technology & Higher Education of Indonesia, and Exploratory Research Grant Scheme (ERGS) RDU130604 & Fundamental Research Grant Scheme (FRGS) RDU140137 from Ministry of Higher Education of Malaysia.

REFERENCES

[1] M. Li, L. K. Kwoh, C. Yang, and S. C. Liew, “3D Building Extraction With Semi-Global Matching From Stereo Pair Worldview-2 Satellite Imageries,” in IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2015, pp. 3006–3009.

[2] V. E. Schmidt and Y. Rzhanov, “Measurement of micro-bathymetry with a GOPRO underwater stereo camera pair,” in OCEANS 2012 MTS/IEEE: Harnessing the Power of the Ocean, 2012, pp. 1–6.

[3] M. Malesa, K. Malowany, U. Tomczak, B. Siwek, M. Kujawińska, and A. Siemińska-Lewandowska, “Application of 3D digital image correlation in maintenance and process control in industry,” Comput. Ind., vol. 64, no. 9, pp. 1301–1315, 2013.

[4] R. Lagisetty, N. K. Philip, R. Padhi, and M. S. Bhat, “Object detection and obstacle avoidance for mobile robot using stereo camera,” in IEEE International Conference on Control Applications (CCA), 2013, pp. 605–610.

[5] A. Hasni, A. Hasan, R. A. Hamzah, and M. H. Johar, “Region of Interest in Disparity Mapping for Navigation of Stereo Vision Autonomous Guided Vehicle,” in International Conference on Computer Technology and Development, 2009, pp. 98–102.

[6] M. Bosse and R. Zlot, “Continuous 3D scan-matching with a spinning 2D laser,” 2009 IEEE Int. Conf. Robot. Autom., pp. 1–8, 2009.

[7] Y. Yu, J. Li, H. Guan, F. Jia, and C. Wang, “Three-Dimensional Object Matching in Mobile Laser Scanning Point Clouds,” IEEE Geosci. Remote Sens. Lett., vol. 12, no. 3, pp. 492–496, 2015.

[8] S. Wibirama and R. Hidayat, “Rekonstruksi dan Visualisasi Objek 3-D Berbasis Algoritma Direct Linear Transformation,” J. Nas. Tek. Elektro Teknol. Inf., vol. 1, no. 2, pp. 32–39, 2012.

[9] R. B. Rusu, Z. C. Marton, N. Blodow, M. Dolha, and M. Beetz, “Towards 3D Point cloud based object maps for household environments,” Rob. Auton. Syst., vol. 56, no. 11, pp. 927–941, 2008.

[10] S. Yamao, M. Miura, S. Sakai, K. Ito, and T. Aoki, “A Sequential Online 3D Reconstruction System Using Dense Stereo Matching,” in 2015 IEEE Winter Conference on Applications of Computer Vision, 2015, pp. 341–348.

[11] G. Heiko Hirschmüller (Inst. of Robotics & Mechatronics Oberpfaffenhofen, German Aerosp. Center, Wessling, “Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information,” in IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2005, pp. 807 – 814.

[12] Z. Zhang, “A Flexible New Technique for Camera Calibration (Technical Report),” IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 11, pp. 1330–1334, 2002.

[13] J. Heikkila and O. Silven, “A Four-step Camera Calibration Procedure with Implicit Image Correction,” in Proceedings - IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1997, pp. 1106–1112.

[14] G. Egnal, “Mutual Information as a Stereo Correspondence Measure,” Tech. Reports, Dep. Comput. Inf. Sci. Univ. Pennsylvania, no. MS-CIS-00–20, 2000.

[15] J. Kim, V. Kolmogorov, and R. Zabih, “Visual Correspondence Using Energy Minimization and Mutual Information,” in Proceedings of the 9th IEEE International Conference on Computer Vision, 2003, pp. 1033–1040.