EMITTER International Journal of Engineering Technology
Vol. 10, No. 2, December 2022, pp. 294~310
DOI : 10.24003/emitter.v10i2.700

Omnidirectional Stereo Vision Study from Vertical and Horizontal Stereo Configuration

Husein Aji Pratama[1,2], Bima Sena Bayu Dewantara[1,2], Dadet Pramadihanto[1,2]

[1]Department of Informatics and Computer Engineering, Politeknik Elektronika Negeri Surabaya, Indonesia
[2]Robotics and Intelligent System Center (RoISC), Politeknik Elektronika Negeri Surabaya, Indonesia
Correspondence Author: dadet@pens.ac.id

Received September 2, 2022; Revised October 4, 2022; Accepted November 5, 2022

Abstract

In stereo vision, an omnidirectional camera has high distortion compared to a standard camera, so the camera calibration is very decisive in its stereo matching. In this study, we will perform stereo matching for an omnidirectional camera with vertical and horizontal configuration so that the result of the image's depth has a 360-degree field of view by transforming the image using a calibration-based method. The result is that by using a vertical camera configuration, the image can be stereo matched directly, but by configuring a horizontal image, it is necessary to carry out a different stereo-matching process in each direction. Stereo matching with the semi-global matching method has better image results than block matching with more image objects detectable by the semi-global block matching method with a maximum disparity value of 32 pixels and with a window size of 21 pixels.

Keywords: Stereo camera, Omnidirectional, Stereo matching, Stereo calibration, Omnidirectional calibration

1. INTRODUCTION

Since their establishment, research conducted by RoISC (previously called ER2C) has progressed rapidly and has become pervasive in various applications. RoISC (Robotics and Intelligent System Center) technology has been penetrated into a wider area in various research fields, ranging from humanoid robots, to computer vision [1]. In computer vision research fields, various studies such as robust human body orientation estimation [2], density-based clustering for stacked pipe object [3], and so on have been carried out recently. Not content with that, RoISC is currently preparing to delve deeper into computer vision research and plans to achieve novelty by combining stereo vision features with omnidirectional cameras features. Those domains are novel concept and seldom explored by the researchers. The knowledge gained from previous researches would be very useful for RoISC to preparing for the new research, and this process would continue indefinitely.

The development of stereo vision has been widely used in robotics, attracted the interest of RoISC to research the field. Stereo vision allows the robots to perceive the depths of points on an object in 3D scenery, working by capturing two images of the same scene with two distinct cameras placed at a specific distance apart and from two different viewpoints [4]. Previously, RoISC has discussed such topics, as presented by Dadet et al. [5], which shows a method for reducing the error rate of vacant space in the depth data by combining a stereo camera and structure sensor. Other than that, numerous potential uses have sparked the development of stereo vision, ranging from vehicle safety [6,7], distance measurement [8], autonomous underwater vehicle [9,10], to the 3D head pose estimation [11]. However, as stereo vision cameras typically only shoot in one direction, their field of view (FOV) limited.

Compared to ordinary cameras, omnidirectional cameras have a wider horizontal field of view (FOV). An omnidirectional vision system offers a 360-degree panorama in the horizontal direction and it consists of a mirror and CCD, making the number of cameras required and total expenses are decreased [12]. Therefore, omnidirectional camera has gained popularity among researchers in various domains, such as video surveillance [13–15], robot driving [16,17], mobile robot competition [18,19],and augmented reality [20,21]. However, only a few studies have performed a stereo vision based on omnidirectional cameras because they have a more difficult on calibration system and more complex on stereo matching process.

This study aims to solve the problem between stereo vision and omnidirectional cameras, by reports a research in the field of stereo computer vision using an omnidirectional camera. In contrast to other studies that use a single-direction camera to perform stereo matching, this study will use an omnidirectional camera, so the camera calibration process will be very decisive in the stereo matching process. In addition, we have already known that stereo matching for omnidirectional cameras is often done with a vertical camera configuration as in the study [22], while the vertical and horizontal configurations have their respective advantages and disadvantages as described in the study [23]. This study carried out research for omnidirectional stereo matching with horizontal configuration with slicing-stitching and vertical techniques and then compared the two of these techniques and camera calibration process.

2. RELATED WORKS

The applications of stereo vision in the field of robotics such as mapping [24] and obstacle avoidance [25] are strongly influenced by the performance of processing images into three dimensions and sensing the distance of an object. In stereo vision, especially in omnidirectional cameras, camera calibration plays an important role in the stereo-matching process. Omnidirectional cameras have greater distortion than ordinary cameras so the calibration process will be more difficult than a standard camera.

Discussing the omnidirectional camera calibration, research on [26] carried out the omnidirectional camera calibration from the extrinsic and intrinsic camera parameters obtained from a chessboard pattern, then the program was optimized and a toolbox was made from MATLAB. Developing from there, Gong et al. [27] aimed to improve the precision of omnidirectional image feature point extraction to enhance calibration accuracy by proposing high-precision feature point extraction based on projection. The output of these research will produce an omnidirectional image that has been transformed into a rectangular image (panoramic image). A more recent study, research [28] perform the transformation of an omnidirectional image into a panorama using a geometric approach. However, the resulting image from the transformation has image damage at the bottom because the pixel position of the inner circle follows the position of the outer circle, so it is necessary to set the parameters and adjust the image size. In the study used by research [29], they showed differences in stereo cameras with vertical and horizontal configurations where stereo camera with vertical configurations had weaknesses in detecting vertical objects, while stereo camera with horizontal configurations had weaknesses in detecting horizontal objects. As a result of this research, they were able to combine vertical and horizontal disparity images from two camera settings (vertical and horizontal) that captured the same image object.

3. ORIGINALITY

Stereo cameras have been commonly used in various studies, but these studies were using standard camera (non-catadioptric mirror or non-hyperbolic lens) to doing the stereo vision, in the other hand stereo vision with omnidirectional cameras are commonly used with vertical configuration because there is a lot easier than horizontal configuration. So that the authors want to compare stereo matching with vertical and horizontal configurations using an omnidirectional camera. By using an omnidirectional camera, the calibration process will greatly determine how good the depth of the stereo camera is. The differences between previous studies are authors use omnidirectional camera instead of standard camera and in addition, authors want to develop a stereo matching technique with a horizontal configuration for omnidirectional cameras where there has been no research on this configuration before.

4. System Design

4.1. Omnidirectional Camera System

From research by Scaramuzza, et al in [26] and [30] they generalized a single-view angle omnidirectional camera model that is dioptric or catadioptric. They identified two different references namely the camera image plane (u', v') and the camera sensor plane $(u", v")$. In the case of a catadioptric camera, the different planes are shown in Figure. 1. All coordinates will be expressed in a coordinate system placed at O with the z-axis parallel to the sensor axis.

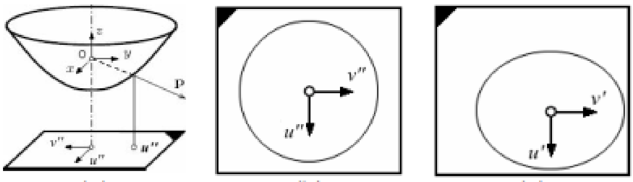


Figure. 1. Omnidirectional coordinate system (a), *Sensor plane* with matrix coordinate(b), *Camera image plane* with pixel coordinate that depends on affine transformation.

If X is the scene point, then assume that $u"= [u", v"]^T$ to be the projection of X on the sensor plane, and the $u'=[u', v']^T$ is the projection of camera field. They also introduced an image projection function which captures the relationship between point $u"$ in the sensor plane, and a vector p that radiates from point of view O to scene point X. By applying the above function, the overall mathematical model of an omnidirectional camera is

$$\lambda. p = \lambda. g(u") = \lambda. g(Au' + t) = PX, \qquad \lambda > 0 \tag{1}$$

Where X represents homogeneous coordinates and P represents the projection matrix. By calibrating the camera, it means that we estimate the matrix A, t and the non-linear function g so that all vectors $g(Au' + t)$ satisfy the equation (1). Assume that g has the following equation

$$g(u,\mathrm{v}) = (u,\mathrm{v}, f(u", v"))^T \tag{2}$$

Where f is the symmetrical rotation about the sensor axis. Since in this case we are using an omnidirectional camera, it is assumed that the mirror is perfectly symmetrical about each axis. So, the equation (1) can be written as

$$\lambda. \begin{bmatrix} u" \\ v" \\ w" \end{bmatrix} = \lambda. g(Au' + t) = \lambda. \begin{bmatrix} (Au' + t) \\ f(u", v") \end{bmatrix} = P \cdot X, \qquad \lambda > 0 \tag{3}$$

4.2. Calibration-based Image Transformation

The essence of omnidirectional camera calibration is estimating parameters$[A, t, a_0, a_1, \dots, a_N]$. To estimate parameter A dan t research [26] was using iterative method without requiring the visibility of the outer circle boundary. First of all, assume that $u" = a. u'$ So by substituting the relationship between equations (2) and (3), the following equation will be produced

$$\lambda. \begin{bmatrix} u" \\ v" \\ w" \end{bmatrix} = \lambda. g(a. u') = \lambda. \begin{bmatrix} a. u' \\ a. v' \\ f(a. \rho') \end{bmatrix} = \lambda. a \begin{bmatrix} u' \\ v' \\ a_0 + \dots + a_N p'^N \end{bmatrix} = PX, a > 0 \tag{4}$$

u' and v' are the pixel coordinates of the point measured from the center of the image, and p' is the Euclidean distance a can be directly integrated into the depth factor. So that in omnidirectional camera calibration only N parameters $(a_0 + \dots + a_N)$ need to be estimated [26]. In the calibration procedure that takes pictures of the chessboard pattern, we do not know the geometric position with respect to the camera. This is related to the camera's extrinsic parameters, namely matrix $R = [r_1, r_2, r_3]$ and translation t. Assume that I^i is the pattern images used for calibration. $M_{ij} = [X_{ij}, Y_{ij}, Z_{ij}]$ are the 3D coordinates of the image pattern, and $m_{ij} = [u_{ij}, v_{ij}]$ are the pixel coordinates in the image plane. We know that if the image pattern is a planar plane, we have$Z_{ij} = 0$, so the equation (4) becomes

$$\lambda_{ij}.P_{ij} = \lambda \cdot \begin{bmatrix} u_{ij} \\ v_{ij} \\ a_0 + \cdots + a_N {\rho_{ij}}^N \end{bmatrix} = P^i \cdot X = \begin{bmatrix} r_1^i \, r_2^i \, r_3^i \, t^i \end{bmatrix} \cdot \begin{bmatrix} X_{ij} \\ Y_{ij} \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} r_1^j \, r_2^j \, t^j \end{bmatrix} \begin{bmatrix} X_{ij} \\ Y_{ij} \\ 1 \end{bmatrix} \quad (5)$$

4.2.1. Extrinsic Parameters Estimation

To estimate the extrinsic parameters, they eliminated from the depth scale λ_{ij} so that each point P_j in the pattern taken has three homogeneous equations below

$$v_j \cdot (r_{31}X_j + r_{32}Y_j + t_3) - f(\rho_j) \cdot (r_{21}X_j + r_{22}Y_j + t_2) = 0 \quad (6.1)$$

$$f(\rho_j) \cdot (r_{11}X_j + r_{12}Y_j + t_1) - u_j \cdot (r_{31}X_j + r_{32}Y_j + t_3) = 0 \quad (6.2)$$

$$u_j \cdot (r_{21}X_j + r_{22}Y_j + t_2) - v_j \cdot (r_{11}X_j + r_{12}Y_j + t_1) = 0 \quad (6.3)$$

X_j, Y_j and Z_j has known, so are u_j, v_j. So, by entering the unknown variable $r_{11}, r_{12}, r_{21}, r_{22,}t_1, t_2$ into a vector, we can get the calibration pattern as a system of linear equation below

$$M \cdot H = 0, \quad (7)$$

Where

$$H = [r_{11}, r_{12}, r_{21}, r_{22,} t_1, t_2]^T,$$
$$M = \begin{bmatrix} -v_1X_1 & -v_1Y_1 & -u_1X_1 & u_1Y_1 & -v_1 & u_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -v_LX_L & -v_LY_L & -u_LX_L & u_LY_L & -v_L & u_L \end{bmatrix} \quad (8)$$

The linear estimation of H can be obtained by minimizing the least squares criterion by using Singular Value Decomposition (SVD).

4.2.2. Intrinsic Parameters Estimation

In the previous step, it was explained how to find extrinsic parameters on the camera using equation (6.3). In the process of estimating the intrinsic parameters, it is explained in research [26] that we need to substitute the estimated values in the equation (6.1) and (6.2) so that the intrinsic parameters $a_0, a_1, \ldots a_N$ can be determined. The estimated equation from the observation of several image patterns is

$$\begin{bmatrix} A_1 & A_1{\rho_1}^2 \dots & A_1{\rho_1}^N & -v_1 & 0 & \dots & 0 \\ C_1 & C_1{\rho_1}^2 \dots & C_1{\rho_1}^N & -u_1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ A_K & A_K{\rho_K}^2 \dots & A_K{\rho_K}^N & 0 & 0 & \dots & -v_K \\ C_K & C_K{\rho_K}^2 \dots & C_K{\rho_K}^N & 0 & 0 & \dots & -u_K \end{bmatrix} \begin{bmatrix} a_0 \\ a_2 \\ \vdots \\ a_N \\ t_3^1 \\ t_3^2 \\ \vdots \\ t_3^K \end{bmatrix} = \begin{bmatrix} B_0 \\ D_1 \\ \vdots \\ B_K \\ D_K \end{bmatrix} \tag{9}$$

Where

$$A_i = r_{21}^i X^i + r_{22}^i X^i + t_2^i \,, B_i = v_i \cdot \left(r_{31}^i X^i + r_{32}^i Y^i\right), C_i = r_{11}^i X^i + r_{12}^i Y^i + t_1^i \,,$$

$$D_i = u^i \cdot \left(r_{31}^i X^i + r_{32}^i X^i\right)$$

The solution of least squares is obtained by using pseudo inverse. Thus, the intrinsic parameters of the camera $a_0, a_1, \dots, a_{n,}$ can be determined.

4.3. Geometrical Image Transformation

The method used for this transformation is remapping each pixel of the omnidirectional image or it can be called unwrapping. To be able to transform the epipolar image into a panoramic form, research by [28] brings up how the algorithm in image transformation is performed.

The first step that needs to be done is to find the coordinates of each pixel in the epipolar image. Then map the coordinates in the epipolar x, and y coordinates into the p dan θ coordinates in the panoramic image. To map the image using the Pythagorean formula to calculate the value of r in an omnidirectional image with the following equation.

$$r = \sqrt{(X_{omni} - X_{center})^2 + (Y_{omni} - Y_{center})^2} \tag{10}$$

To obtain the θ parameter on the omnidirectional image, we can use the inverse of tan θ with the following equation.

$$\theta = tan^{-1} \frac{Y_{omni} - Y_{center}}{X_{omni} - X_{center}} \tag{11}$$

So that new coordinates are found where the coordinates in the panoramic image are a dan β if each has a value of 1 and 0.5 will produce the maximum resolution.

$$(y_{unwrap}, x_{unwrap}) = (a(Rmax - p), \beta(Rmax.\theta)) \tag{12}$$

a and β are the a scaling constant that affects the resolution and ratio of the output image.

4.4. Stereo Calibration

We performed stereo calibration by manually observing with the human eye by looking at the tangent lines. For vertical stereo calibration, we generate multiple vertical lines on each image at the same pixel position. Figure. 2 is the result of vertical camera calibration.

Figure. 2. Vertical stereo calibration.

It can be seen that in each image, for vertical stereo camera calibration, the previously generated vertical lines hit the same object. This indicates that the stereo calibration is good. As for the horizontal stereo calibration, we generate horizontal lines with the same position for each image as shown in Figure. 3.

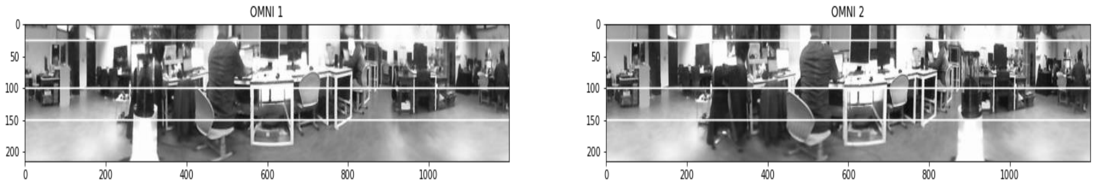


Figure. 3. Horizontal stereo calibration

From Figure. 3, we can see that the same line has the same object behind it such as the neck object, the end of the chair, the end of the table, and so on. So, the camera has been stereo calibrated quite well with a tolerance of a few pixels. Next, we perform the vertical and horizontal stereo matching process on several different image data.

4.5. Horizontal Omnidirectional Perspective

Basically, the stereo matching calculation is performed with the right and left camera perspectives looking at the same object in front of them, which means that there is only one forward-looking perspective from the usual stereo matching calculation. But on the omnidirectional camera there are two different perspective, we can say front-perspective and rear-perspective. Therefore, the application of the common stereo matching method cannot be performed using omnidirectional camera horizontal configuration. Omnidirectional camera that has two perspective illustrated in Figure. 4.

The first perspective is when the two cameras are facing the front or the stereo matching process is in the image with the object in front of the camera. The second perspective is when both cameras perform stereo matching images with objects behind the camera in perspective. The first and second perspectives on the omnidirectional image after undergoing transformation can be seen in Figure. 4.

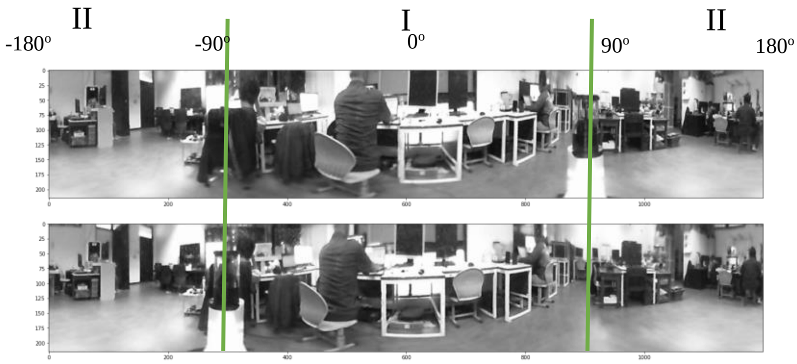


Figure. 4. Perspective field on horizontal stereo omnidirectional camera.

With this perspective, the mathematical equation of horizontal stereo matching turns into

$$\begin{cases} SAD = \sum_{k,l} |g[k,l] - f[x-k,y]| & if(-90^o < image < 90^o) \\ SAD = \sum_{k,l} |g[k,l] - f[x+k,y]| & if(90^o \leq image \geq -90^o) \end{cases} \quad (13)$$

If we want to do stereo matching at -90 to 90 degree of an image, the matching point will be movie to the left. On the other hand, if we want to do stereo matching at 90 increments up to -90, the matching point will be move to the right.

5. EXPERIMENT AND ANALYSIS

5.1. Omnidirectional Vision System

In this study, we will use a catadioptric omnidirectional camera with a diagonal configuration to accommodate the weakness of vertical and horizontal disparities.

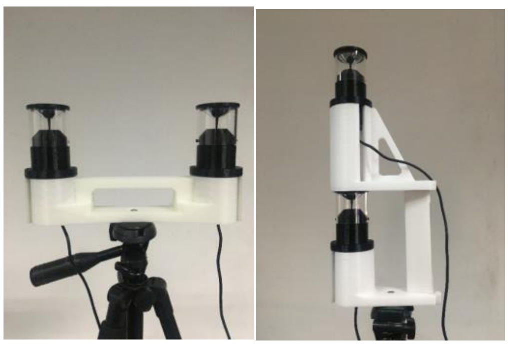


Figure. 5. Horizontal camera design (a) and vertical design(b).

The camera design is made from a 3D printer made of PLA plastic with each arm supporting the camera from a carbon fulcrum. The length of each camera axis measured from the same point between the two cameras is 17.5 cm for the horizontal length and the vertical length. The camera specifications are as follows:

- Resolution is 1280 x 720
- The focal distance is variable
- The interface is USB 2.0

- The ratio of the video stream is 30 fps

5.2. Geometrical Image Transformation

In the Figure. 6 and Figure. 7 the transformation from an omnidirectional image into a panoramic image has been successful but with parameters a = 1, β = 0.5 which results in the maximum image size [1046 x 248], resulting in a broken image at the bottom.

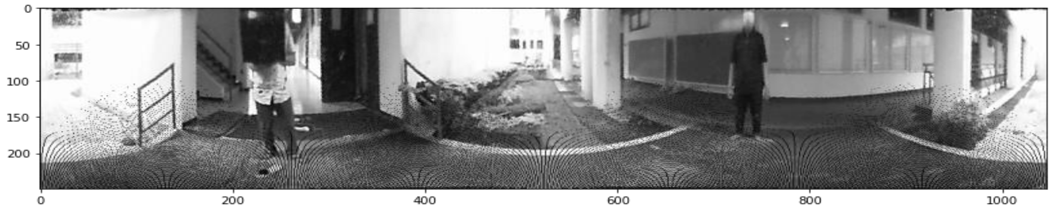


Figure. 6. Result of omnidirectional image transformation, a = 1, β = 0.5

Figure. 7. Result of omnidirectional image transformation, a = 0.6, β = 0.3

This is caused by the length of the inner edge of the circle (minimum circumference of the circle) which must match the length of the circumference of the outermost circle (maximum circumference of the circle) resulting in some pixels not converging at the bottom of the image. The solution is to use an interpolation technique using the median filter technique, or by reducing the image size. In Figure J.3 it is done by reducing the image size with parameters a = 0.6, β = 0.3 so as to produce an image with a size of [628 x 149].

5.3. Calibration-based Image Transformation

The purpose of image rectification is to correct the image from distortion of each frame of the stereo camera pair that is used to align the image in the horizontal and vertical planes. In camera calibration, the chessboard image is used as a reference for doing the image calibration.

We use the calibration algorithm proposed by [30] and described in section 1.2. For each camera, we perform the same calibration so as to get the desired camera parameters. We used 10 images with a chessboard pattern. We used a chessboard which has a 9x6 grid pattern and each square measure 30 mm. The calibration results are shown in table 1.

Table 1. Calibration Parameters

Significance	**Variable**	**Cam 1**	**Cam 2**	**MU**
Image Center	row column	337.017191 331.108538	351.900362 354.315109	pixel
Affine Parameters	c d e	1.000101 -0.000591 -0.000097	1.001649 0.000933 -0.000276	-
Image Size	height	690	690	pixel

	width	689	687	
Taylor Polynomial	a_0 a_1 a_2 a_3 a_4	-1.214194214367139 0 0.000021123140569 -0.000000012161234 -0.000000000001348	-1.233963499140422 0 0.000023665757087 -0.000000034027122 0.000000000043727	-
Reprojection Error	$\sqrt{e_x^2 + e_y^2}$	0.643989	0.627813	pixel

We used Scaramuzza's omnidirectional camera calibration toolbox [30] to perform the calibration. Although this method is more general and can be applied for different types of omnidirectional camera, we has slightly higher average reprojection error that obtained around 0.6 pixels. We capture omnidirectional image at RoISC lab and the result of calibrated omnidirectional image is shown in Figure. 8.

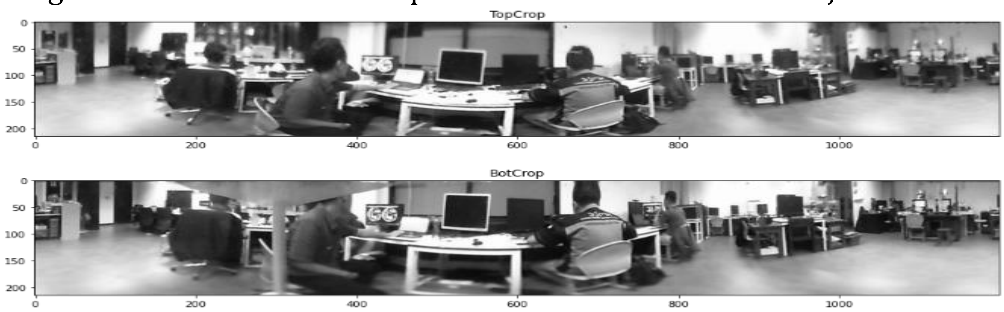
Cam 1	Calibration Result
	
Cam 2	Calibration Result
	

Figure. 8. Calibration Result

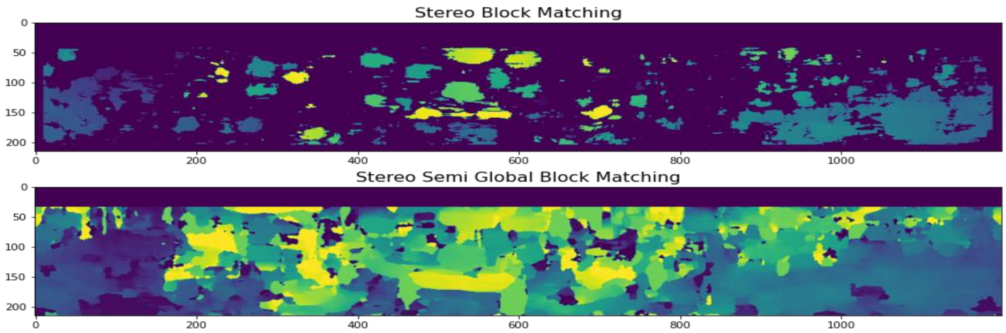
In Figure. 8 columns Cam1 and cam2 are the original omnidirectional images and the column calibration result is where the calibration image is generated. The transformation was quite well produces in the panoramic image but on the top and bottom of panoramic images, there are not perfect horizontally and there is still a few curve. Other than that, when it is compared to the geometry image transformation which has only an image size (628 x 149) to get an image with minimal damage, the calibration-based image transformation has a much larger image size (1200 x 400) and has no image damage due to the transformation process at all so we can conclude that the calibration-based transformation has better result than geometrical-based transformation. So that we decided to conduct our experiments further with calibration-based image transformation model.

5.4. Vertical Stereo Matching

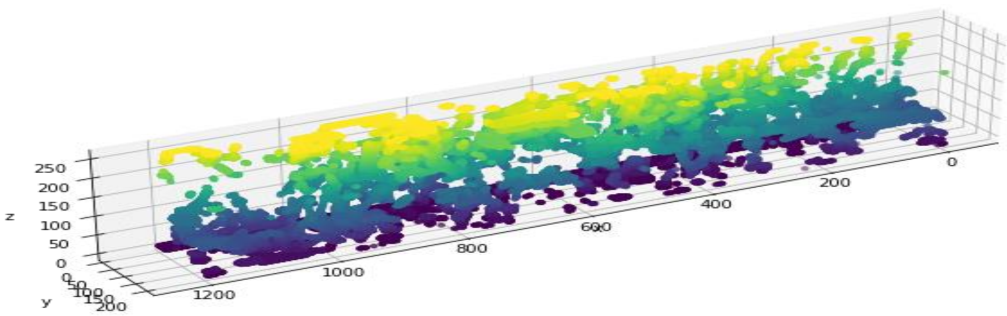
In the vertical stereo matching process, it will look for the difference in the distance (disparity) of the appropriate feature in the upper and lower image pixels. The result of vertical stereo matching is in the form of a disparity map which shows that the darker the image, the farther the distance. On the other hand, the brighter it is, the closer the object is to the camera. For stereo matching, we used a maximum disparity of 32 pixels and a window size of 21 pixels. The following is a complete data result of vertical stereo matching obtained from the ROISC PENS lab. We can also see the 3D plot for the disparity image that shows that color represents the distance of the object.



(a)



(b)



(c)

Figure. 9. The results of vertical stereo matching from top and bottom camera (a), using the BM and SGBM (b), and shown with 3D visualization (c).

Figure. 9 is the result of vertical stereo matching using block-matching and semi-global block matching methods. It can be seen that the vertical stereo matching experiment was carried out from two different perspectives. The performance produced by stereo matching with the block matching method is not so good with many objects that are not matched in both cameras, which is indicated by the production of some parts of the image (in the middle of the image) which are black. It appears that stereo matching using the semi-global block matching method gives better results. We can easily recognize and compare the results of stereo-matching images with the original image for far and near objects.

5.5. Horizontal Stereo Matching

Horizontal stereo matching experiments were also carried out on two different camera orientations and can be seen in Figure. 10. The solution to the problem in the previous section which explains how horizontal omnidirectional cameras cannot be matched directly and is also implemented in horizontal stereo matching by slicing the image section into two (front and back) to display the same left-right image perspective. After the image slicing, we can use general stereo matching function for every pair of front-side and back-side images. Then after all the process is done, the disparity results are combined back into full stereo matching as shown in Figure. 10, full stereo matching. We can also see the 3D plot for the disparity image that shows that color represents the distance of the object.

For this horizontal stereo matching, we used a maximum disparity value of 32 pixels and a window size of 21 pixels. We can see from the results in Figure. 10, some objects have been matched well and can be identified objects far and near quite well. When compared with the block matching method, the semi-global block matching method has better results seen from the smoother disparity results. The results of stereo matching with the block matching method have some parts of the object that are not matched properly so that for the final result on the omni stitching column we use the semi global block matching method.

Source 1

Omni slicing

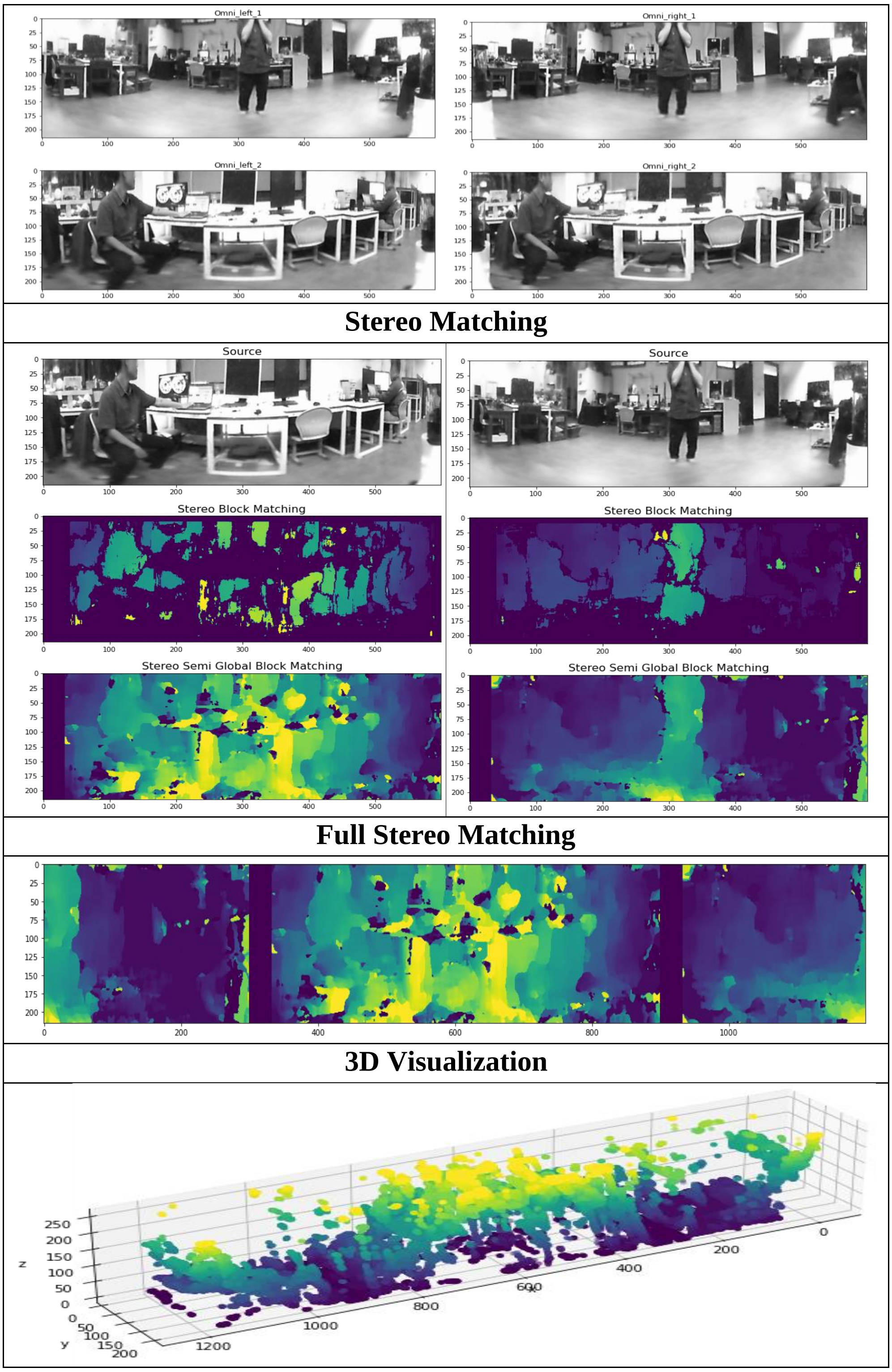


Figure. 10. Horizontal Stereo Matching.

5.6. Stereo Matching Analysis

Vertical and horizontal stereo matching has been described in sections 5.4 and 5.5 so that a comparison of the two methods can be observed. The experiment resulted in the conclusion that the semi-global block matching stereo matching method produces a better disparity image. Next, we carefully analyze the difference between stereo matching with a stereo camera with a vertical combination and a stereo camera with a horizontal combination. We tried to place the camera in the same position and capture the image as in Figure. 11.

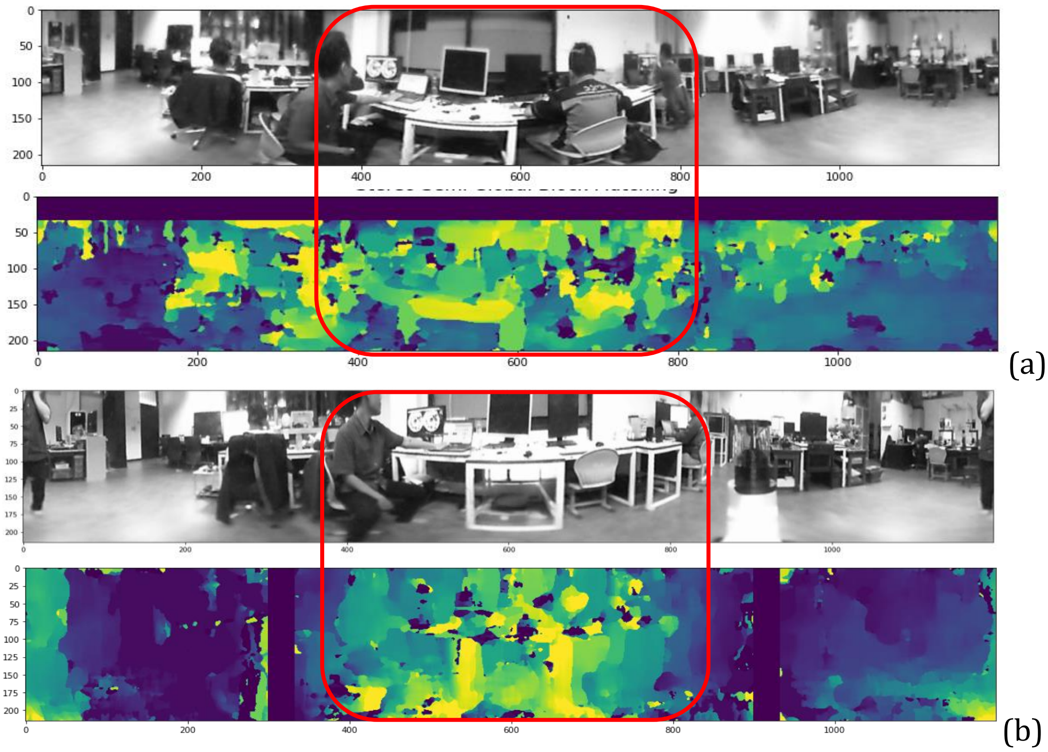


Figure. 11. Vertical stereo camera observation area (a). Horizontal stereo camera observation area (b).

The most noticeable difference is in the center of the image which shows several vertical and horizontal objects firmly such as tables and monitors. From the results of the disparity image, it can be seen that in the vertical camera setting, only the top of the table is detected (which is in the form of a horizontal object) and only a few table legs are detected (the table leg is in the form of a vertical object). Meanwhile, in the horizontal table setting, only the legs of the table (vertical objects) are detected and it is not good at detecting the top of the table (horizontal objects). This study proof same result with research [29] which the vertical stereo camera is better if it detects a horizontal object and a horizontal stereo camera is better if it detects a vertical object is true. For other objects in the image we cannot analyze further because there are no firm horizontal/vertical objects.

6. CONCLUSION

In this study, we conducted research for an omnidirectional camera. Starting from camera calibration to transform omnidirectional cameras into panoramic images, the results of using calibration-based transformation produce better images than geometrical-based ones. Then we perform a stereo calibration which is done manually by using the guide lines placed on the image background to get the same image object along the guide lines. Finally, we performed vertical stereo matching and horizontal stereo matching using the block matching and semi global block matching methods, which resulted in the semi global block matching method providing better disparity image results marked by more black parts of the image and contrasting with other parts, indicating that the section did not match properly. In the horizontal stereo matching process, we consider the horizontal perspective of an omnidirectional camera. When compared to stereo matching with a vertical configuration, the horizontal camera configuration can detect vertical objects well and vice versa, the vertical camera configuration can detect horizontal objects well.

Acknowledgment

Thank you to the Ministry of Education and Culture of the Republic of Indonesia along with the Robotic and Intelligent Systems Center (RoISC) Laboratory and the Politeknik Elektronika Negeri Surabaya (PENS) for the financial and non-financial support that has been given so that this research can be carried out properly.

REFERENCES

[1] Putra VCPH, Apriandy KI, Pramadihanto D, Barakbah AR, **FLoW-Vision: Object Recognition and Pose Estimation System based on Three-Dimensional (3D) Computer Vision**, *2021 International Electronics Symposium (IES)*, pp. 593–9, 2021.

[2] Dewantara BSB, Saputra RWA, Pramadihanto D, **Estimating human body orientation from image depth data and its implementation**, *Machine Vision and Applications*, vol. 33, no. 3, pp. 38, 2022.

[3] Pratama AR, Dewantara BSB, Sari DM, Pramadihanto D, **Density-based Clustering for 3D Stacked Pipe Object Recognition using Directly-given Point Cloud Data on Convolutional Neural Network**, *EMITTER International Journal of Engineering Technology*. vol. 10, no. 1, pp. 153-169, 2022.

[4] Islam A, Asikuzzaman Md, Khyam MO, Noor-A-Rahim Md, Pickering MR, **Stereo Vision-Based 3D Positioning and Tracking**, *IEEE Access*, vol. 8, pp. 138771–87, 2020.

[5] Pramadihanto D, Alfarouq A, Waskitho SA, Sukaridhoto S, **Merging of Depth Image Between Stereo Camera and Structure Sensor on Robot "FloW" Vision**, *International Journal on Advanced Science, Engineering and Information Technology*, Vol. 7, no. 3, pp. 1014, 2017.

[6] Shakeri A, Moshiri B, Garakani HG, **Pedestrian Detection Using Image Fusion and Stereo Vision in Autonomous Vehicles**, *2018 9th International Symposium on Telecommunications (IST)*, p. 592–6, 2018.

[7] Song W, Yang Y, Fu M, Li Y, Wang M, **Lane Detection and Classification for Forward Collision Warning System Based on Stereo Vision**. *IEEE Sensors J*. vol. 18, no. 12, pp. 5151–63, 2018.

[8] Rodríguez-Quiñonez JC, Sergiyenko O, Flores-Fuentes W, Rivas-lopez M, Hernandez-Balbuena D, Rascón R, et al, **Improve a 3D distance measurement accuracy in stereo vision systems using optimization methods' approach**, *Opto-Electronics Review*. vol. 25, no. 1, pp. 24–32, 2017.

[9] Hong S, Chung D, Kim J, Kim Y, Kim A, Yoon HK, **In-water visual ship hull inspection using a hover-capable underwater vehicle with stereo vision**. *J Field Robotics*, vol. 36, no. 3, pp. 531–46, 2019.

[10] Lodi Rizzini D, Kallasi F, Aleotti J, Oleari F, Caselli S, **Integration of a stereo vision system into an autonomous underwater vehicle for pipe manipulation tasks**, *Computers & Electrical Engineering*, vol. 58, pp. 560–71, 2017.

[11] AlQahtani F, Banks J, Chandran V, Zhang J, **Three-Dimensional Head Pose Estimation Using a Stereo Camera Arrangement**, *Proceedings of the International Conference on Machine Vision and Applications - ICMVA*, p. 28–35, 2018.

[12] Gong X, Lv Y, Xu X, Wang Y, Li M, **Pose Estimation of Omnidirectional Camera with Improved EPnP Algorithm**, *Sensors*, vol. 21, no. 12, 2021.

[13] Morbidi F, Caron G, **Phase Correlation for Dense Visual Compass From Omnidirectional Camera-Robot Images**, *IEEE Robot Autom Lett,* vol. 2, no. 2, pp. 688–95, 2017.

[14] Karaimer HC, Baris I, Bastanlar Y, **Detection and classification of vehicles from omnidirectional videos using multiple silhouettes**. *Pattern Anal Applic*, vol. 20, no. 3, pp. 893–905, 2017.

[15] Dias Pais G, Dias TJ, Nascimento JC, Miraldo P, **OmniDRL: Robust Pedestrian Detection using Deep Reinforcement Learning on Omnidirectional Cameras**, *2019 International Conference on Robotics and Automation (ICRA)*, p. 4782–9, 2019.

[16] Patruno C, Colella R, Nitti M, Renò V, Mosca N, Stella E. **A Vision-Based Odometer for Localization of Omnidirectional Indoor Robots**. *Sensors*. 2020 Feb 6;20(3):875.

[17] Liu, Guo, Feng, Yang, **Accurate and Robust Monocular SLAM with Omnidirectional Cameras**, *Sensors*, vol. 19, no. 20, pp. 4494, 2019.

[18] Santosa SR, Dewantara BSB, Wibowo IK, **Ball Detection Under Variety Lighting using Omnidirectional Camera**, *2019 International Electronics Symposium (IES)*, p. 235–40, 2019.

[19] Nurrohmah EA, Sena Bayu B, Bachtiar MM, Kurnianto Wibowo I, Adryantoro R, **Detecting Features of Middle Size Soccer Field using Omnidirectional Camera for Robot Soccer ERSOW**. *2020 International Conference on Smart Technology and Applications (ICoSTA)*, p. 1–6, 2020.

[20] Xie S, Lai PK, Laganiere R, Lang J, **Effective Convolutional Neural Network Layers in Flow Estimation for Omni-Directional Images**, *2019 International Conference on 3D Vision (3DV)*, p. 671–80, 2019.

[21] Kundu AS, Mazumder O, Dhar A, Lenka PK, Bhaumik S, **Scanning Camera and Augmented Reality Based Localization of Omnidirectional Robot for Indoor Application**, *Procedia Computer Science*, vol. 105, pp. 27–33, 2017.

[22] Gao W, Shen S, **Dual-fisheye omnidirectional stereo**, *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, p. 6715–22, 2017.

[23] Kallwies J, Engler T, Forkel B, Wuensche H-J, **Triple-SGM: Stereo Processing using Semi-Global Matching with Cost Fusion**, *2020 IEEE Winter Conference on Applications of Computer Vision (WACV)*, p. 192–200, 2020.

[24] Chen M, Tang Y, Zou X, Huang Z, Zhou H, Chen S, **3D global mapping of large-scale unstructured orchard integrating eye-in-hand stereo vision and SLAM**, *Computers and Electronics in Agriculture*, vol. 187, 2021.

[25] McGuire K, de Croon G, De Wagter C, Tuyls K, Kappen H, **Efficient Optical Flow and Stereo Vision for Velocity Estimation and Obstacle Avoidance on an Autonomous Pocket Drone**, *IEEE Robot Autom Lett.* vol. 2, no. 2, pp. 1070–6, 2017.

[26] Scaramuzza D, Martinelli A, Siegwart R. **A Flexible Technique for Accurate Omnidirectional Camera Calibration and Structure from Motion**, *Fourth IEEE International Conference on Computer Vision Systems (ICVS'06)*, p. 45–45, 2006.

[27] Gong X, Lv Y, Xu X, Jiang Z, Sun Z, **High-Precision Calibration of Omnidirectional Camera Using an Iterative Method**, *IEEE Access*, vol. 7, 152179–86, 2019.

[28] Lin C-H, Chou W-J, **Acceleration of the transformation from elliptic omnidirectional images to panoramic images using graphic processing units**, *2016 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW)*, p. 1–2, 2016.

[29] Kallwies J, Wuensche H-J, **Effective Combination of Vertical and Horizontal Stereo Vision**, *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, p. 1992–2000, 2018.

[30] Scaramuzza D, Martinelli A, Siegwart R, **A Toolbox for Easily Calibrating Omnidirectional Cameras**, *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, p. 5695–701, 2006.