

Software Calibration for Stereo Camera on Stereo Vision Mobile Robot using Tsai's Method

Rostam Affendi Hamzah and Sani Irwan Md Salim

Abstract - An improvement for image distortion in cameras has been an important issue for as long as users have wanted to reproduce or utilize observed information. In this paper, it described about mobile robot using uncalibrated cameras rig which installed on top of it to work as stereo pair. These cameras have to be calibrated by software using Tsai's method. The block matching algorithm is using Sum of Absolute Differences (SAD) to solve the correspondence problem. The image processing part is implemented in Matlab software. The result is a lot of matching points discovered using a software calibration result from Tsai's method.

Index Terms - Block matching; disparity mapping; epipolar line; stereo vision; uncalibrated camera

I. INTRODUCTION

A camera consists of an image plane which provides a transformation between object space and image space. This transformation cannot be described perfectly by a perspective transformation because of distortions which occur between points on the object and the location of the images of those points. These distortions can be modeled [1]. However, the model may only be an approximation to the real relationship. How closely the model match to reality will depend on the model and how well the model's parameters can be estimated. Choosing parameters which are both necessary and sufficient has taxed those involved in the process of lens calibration for as long as lenses have been used to make precise measurements.

A. Mobile Robot Using Uncalibrated Stereo Pair

Basically, the two stereo cameras on top of mobile robot Figure 1 take pictures of a same object but with different view. The two dimension images on the plane of projection represent the object from camera view. These two images contain some encrypted information, the image-depth of each other. This information is the third dimension of two dimension images. Therefore, the object distance and its depth can be determined by using the stereo cameras [2]. With referring to Figure 2 the distance between two points of view is called the baseline [3]. The baseline's distance end to end affects on the range resolution, which establish the range of depth that can be calculated. A difference between scenes of the same object on the two images is called disparity [4].

Rostam Affendi Hamzah is with Faculty of Electronics & Computer Engineering, UTeM, Melaka, Melaka, Malaysia. (email: rostamaffendi@utem.edu.my)

Sani Irwan Md Salim is with Faculty of Electronics & Computer Engineering, UTeM, Melaka, Melaka, Malaysia. (email: sani@utem.edu.my)

The stereo system uses horizontal baseline. The cameras are placed at the same elevation. The process of stereo vision is



Figure 1. Stereo camera placement

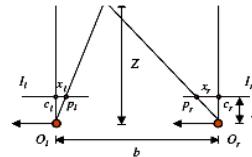


Figure 2. Basic principle of stereo vision

then typically defined as finding a match between features in left and right images [5]. The stereo baseline b is the distance between the centres of the projection O_l and O_r . The x_l and x_r are the coordinates of p_l and p_r with respect to the principal points cl and cr . Where the f is a focal length of the camera. The depth of P is uniquely determined by the following equation [4]:

$$\frac{b+x_l-x_r}{z-f} = \frac{b}{z} \quad (1)$$

And let the disparity $d = x_r - x_l$, then the depth of point P is:

$$Z = f \frac{b}{d} \quad (2)$$

The corresponding two images representing the same point of the scene is called disparity matching. The set of displacements between matched pixels is usually indicated as disparity map [6].

II. TSAI'S METHOD FOR UNCALIBRATED CAMERAS RIG

Construction of a full model from the stereo pair requires calibration of the camera system using software. From a stereo camera system (stereo rig) has accurately specified intrinsic and extrinsic parameters for both cameras. According to [7] the intrinsic camera parameters specify a pinhole camera model with radial distortion. The pinhole model is characterized by its focal length, image centre, pixel spacing in two dimensions and the radial distortion is characterized by a single parameter. The extrinsic parameters describe the relative position and orientation of the two cameras. Intrinsic parameters for a given camera are constant, assuming the physical parameters of the optics do not change over time, and thus may be pre-calculated. Extrinsic parameters depend on the relative camera poses and will be

constant if the cameras are fixed relative to one another [7]. Both intrinsic and extrinsic calibration parameters are calculated using the Tsai's method from a tool by Matlab provided by [8]. There are images of a calibration target, see Figure 3, consisting of images of left and right scene. It is totally uncalibrated images due to unaligned each comparing image left to right. This paper is presented using twenty images (left and right target) for calibration process. Figure 4 shows the result of extrinsic parameters which produces the rotation and translation of the chess board during calibration process from figure 3.

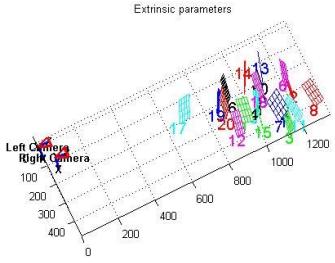


Figure 4. Extrinsic parameters in 3D

The flowchart Figure 5 shows the steps of camera calibration using the Tsai's method in Matlab toolbox. The first step is to get a set of images in digital form and start to evaluate the error between the images of left and right. If the images are not converge each other then the system will adjust the value of the camera evaluation until they converge. The adjusted value or parameters will be used as a result for calibration process to be used in rectifying process [9]. The result contains of intrinsic parameters TABLE I and TABLE II and extrinsic parameters in TABLE III. These values are represented in pixel intensity values. As a result in matlab software, Tsai's method produces the data in extension of .mat files. These files to be called by main program for rectification process to align the epipolar line on stereo images.

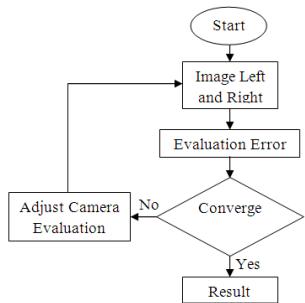


Figure 5. Flowchart of Tsai's method

TABLE I. INTRINSIC PARAMETERS OF LEFT CAMERA

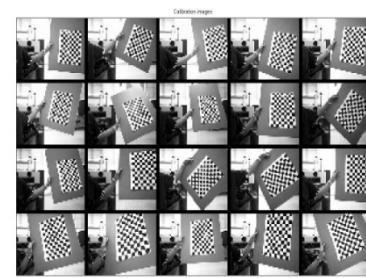
Focal length	[372.90575 373.58482] ± [5.14011 5.07494]
Principal point	[159.76547 91.23921] ± [5.46331 5.76730]
Skew	[0.00000] ± [0.00000]
Distortion	[0.19492 -1.07074 -0.01118 -0.00091 0.00000]

TABLE II. INTRINSIC PARAMETERS OF RIGHT CAMERA

Focal length	[373.83866 373.83178] ± [4.82569 4.85865]
Principal point	[167.07304 139.96128] ± [5.64311 5.82166]
Skew	[0.00000] ± [0.00000]
Distortion	[0.19645 -0.84041 -0.01428 0.00668 0.00000]

TABLE III. EXTRINSIC PARAMETERS (LEFT AND RIGHT CAMERA)

Rotation Vector	[-0.00757 0.16506 -0.00930] ± [0.01161 0.01458 0.00162]
Translation Vector	[-47.82594 1.48281 3.10607] ± [0.77822 0.53278 4.14571]



(a)



(b)

Figure 3. Calibration target on chessboard using Tsai's method

III. IMAGE RECTIFICATION

The rectification of stereo image pair can be carry out under the condition of camera's parameters in Tsai's method. To quickly and accurately search the corresponding points along the scan-lines, rectification of stereo pair are performed so that corresponding epipolar lines are parallel to the horizontal scan-lines and the difference in vertical direction is zero shown by Figure 6. Image rectification is the undistortion according to the calibration parameters calculated in the Tsai's calibration. After all intrinsic and extrinsic camera parameters are calculated, they can be used to rectify images according to the epipolar constraint [10]. The equation below is used to rectify the images in Matlab.

$$I_{new}(x_0, y_0) = a_1 I_{old}(x_1, y_1) + a_2 I_{old}(x_2, y_2) + a_3 I_{old}(x_3, y_3) + a_4 I_{old}(x_4, y_4) \quad (4)$$

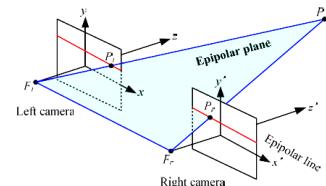


Figure 7. Epipolar plane of stereo vision

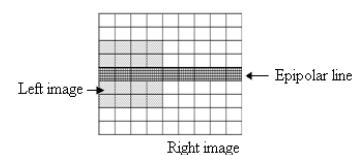


Figure 8. SAD block matching algorithm

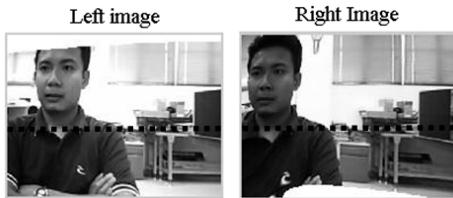


Figure 6. Rectified image using Tsai's method of camera calibration result

With I_{new} and I_{old} as the original and the rectified image and the blending coefficients a_i separate for each camera. Figure 6 is the output of rectified stereo image with a size of 320x240. The horizontal line (dotted line) for both images indicates the left and right image is horizontally aligned.

IV. BLOCK MATCHING ALGORITHM

The sum of absolute differences algorithm is implemented in the image processing software Matlab. The SAD algorithm from the equation below shows an area-based correspondence algorithm [11] [12]. It computes the intensity differences for each center pixel of (i, j) as follows:

$$SAD_v(x, y) = \sum_i \sum_j \|g_v(x+i, y+j) - g_{v-1}(x+v+i, y+j)\| \quad (5)$$

Or in easy interpretation the reference point of the left image minus with the right image at the same epipolar plane Figure 7. It sums up the intensities of all surrounding pixels in the neighborhoods for each pixel in the left image. The absolute difference between this sum and the sum of the pixel, and its surrounding, in the right image is calculated. The minimum over the row in the right image is chosen to be the best matching pixel. The disparity then is calculated as the actual horizontal pixel difference. The plane passing through the camera centers and the object point in the scene is called the epipolar plane Figure 7. The intersection of the epipolar plane with the image plane is called epipolar line. By referring the epipolar geometry, correspondences at point P_l and P_r must lie on the epipolar line [2].

According to [12] the SAD algorithm works on each block from the left image is matched into a block in the right image by shifting the left block over the searching area of pixels in right image as shown in Figure 8. Ideally, for every pixel mask within the original image there should be a single mask within a second image that is nearly identical to the original and thus the SAD for this comparison should be zero [13].

V. DISPARITY MAPPING

Together with the stereo camera parameters from Tsai's calibration and the disparity between corresponding stereo points, the stereo images distances can be retrieved. In order to find corresponding pairs of stereo points, they first have to be compared for different disparities, after which the best matching pairs can be determined. The maximum range at which the stereo vision can be used for detecting obstacles depends on the image and depth resolution [14]. Absolute differences of pixel intensities are used in the algorithm to compute stereo similarities between points. By computing the sum of the absolute differences SAD for pixels in a window surrounding the points, the difference between

similarity values for stereo points can be calculated. The disparity associated with the smallest SAD value is selected as best match [10]. Figure 9 shows the disparity mapping using SAD algorithm that taken from Figure 6. As a comparison of mapping the disparity, Figure 10 shows the output without using a camera calibration result. Even from inspection of the mapping, Figure 9 shows a lot of matching points detected compared to Figure 10.

VI. DEPTH ESTIMATION

Together with the stereo camera parameters from software calibration and the disparity between corresponding stereo points, real-world distances can be retrieved. In order to find corresponding pairs of stereo points, they first have to be compared for different disparities, after which the best matching pairs can be determined. The maximum range at which the stereo vision can be used for detecting obstacles depends on the image and depth resolution [10]. The estimation of the obstacle's range in this paper is using curve fitting tool (cftool) Figure 11 provided by Matlab software to determine the range according to the pixel values.

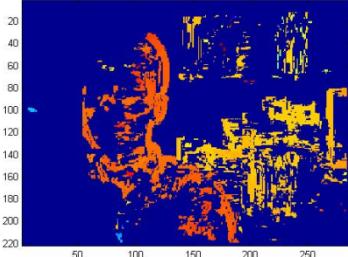


Figure 9. Disparity mapping using a calibration result

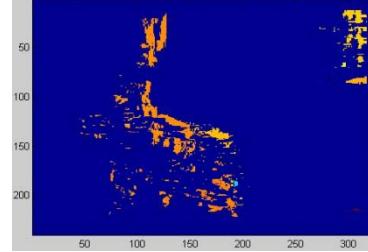


Figure 10. Disparity mapping without using a camera calibration result

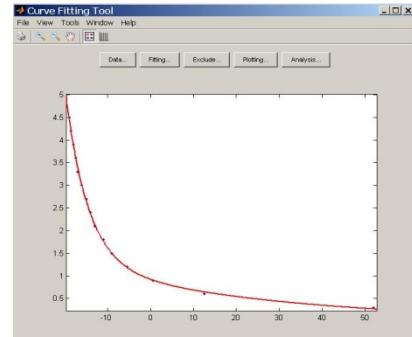


Figure 11. Curve fitting tool window with plotted range versus disparity values

Each pixel in the mapping of disparity will be calculated through the curve fitting tool and the coordinate of horizontal is referring to left image. The equation below shows the distance or depth of obstacles that has been used in this paper.

$$Distance = ae^{bx} + ce^{dx} \quad (6)$$

The value of a, b, c and d is a constant value from curve fitting tool in Matlab. The value of x represents a pixel value in disparity mapping. The calculated distance has an effective range of detection for the mobile robot.

VII. OBSTACLE DETECTION AND CLUSTERING

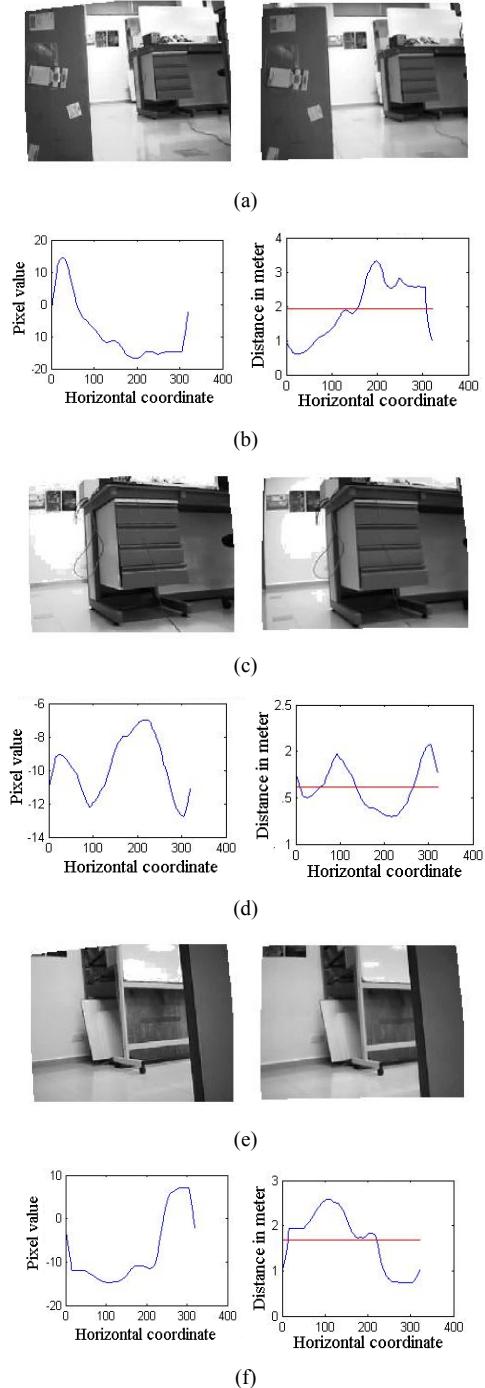


Figure 12. Sample images taken by mobile robot and the range of obstacle detection

In the obstacle detection step, the points belonging to obstacles must be found. For data combination in image processing software, it is important to cluster or group the obstacle points in such a way that individual obstacles can be identified. This makes the mobile robot movement become

easy to navigate as assumption a big object or obstacle in front of it. If there are obstacles at closest range, the stereo program will filter using Gaussian lowpass filter [15](7) to determine the object and cluster the pixels location.

$$H(u, v) = e^{-D^2(u,v)/2\sigma^2} \quad (7)$$

VIII. EXPERIMENT AND RESULT

The result of the disparity values is shown by TABLE IV. The maximum detected range is 4.5 meter and -19 disparity value. And the minimum detected range is about 0.3 meter with 52 disparity value. This region refers to an effective range of mobile robot detection. These readings are reference to the mobile robot movement. Some samples of images taken by mobile robot are presented in Figure 12 and the result of detected obstacle and range in the graph prove that the range estimation by cfotool. These results can be compared with the graphs of pixel value for each detected obstacle.

TABLE IV. RESULT OF DISPARITY VALUE AND RANGE.

Pixel Value	52	12.5	0.5	-5.5	-9	-11	-13	-14	-15	-16	-17	-17.5	-18	-18.5	-19
Range (m)	0.3	0.6	0.9	1.2	1.5	1.8	2.1	2.4	2.7	3	3.3	3.6	3.9	4.2	4.5

The experiment was conducted in Mechatronic Lab for obstacle detection and range estimation by mobile robot. The continuous process of image capturing by mobile robot are presented in Figure 12(a)(c)(e). The result of distance and pixel values are presented in the graphs Figure 12(b)(d)(f). The result could be explained when the pixel value is large, the range is closer to the mobile robot and if the pixel values are small, the range located faraway from mobile robot. So the pixel works contrary to the range shown by Figure 13.

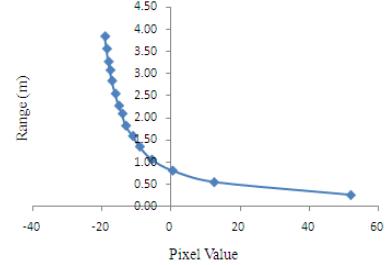


Figure 13. Relation of pixel value and range

IX. CONCLUSION

The result from camera calibration process is significantly important to the rectification process. The uncalibrated stereo rig has been experimentally confirmed showing that an impressive performance if the stereo camera has calibrated using software. A lot of matching points detected by the SAD if compared to original image without go through the software calibration. The mobile robot is able to determine the range of the obstacles by mapping the environment. This is proved by the experiment conducted to determine the range by using the camera calibration result.

REFERENCES

- [1] J. Xu and O. Chutatape, Comparative study of two calibration methods on fundus camera," in Proc. 26th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society, pp. 576-579, (Cancum, Mexico), 2003
- [2] Teerapat Chinapirom, U.W., and Ulrich Rückert, Steroscopic Camera for Autonomous Mini-Robots Applied in KheperaSot League. System and Circuit Technology, Heinz Nixdorf Institute, University of Paderborn, 2001.
- [3] Konolige, K., Stereo Geometry. SRI International. 1999.
- [4] B. Klaus, P.H., Robot Vision (MIT Electrical Engineering and Computer Science Series). MIT Press, McGraw-Hill Book Company, Cambridge, MA, 1986.
- [5] Foggia, P., A New Approach for Stereo Matching in Autonomous Mobile Robot Applications. Dipartimento di Informatica e Sistemistica, Università di Napoli, Via Claudio 21, I80125, Napoli, ITALY, 2007.
- [6] Digney, S.S.a.B., Autonomous Cross-Country Navigation Using Stereo Vision. CMU-RI-TR-99-03, Carnegie Mellon University Technical report, 1999.
- [7] J. Steele, C.D., T. Vincent and M. Whitehorn, 2001. Developing stereovision and 3D modelling for LHD automation. 6th International Symposium on Mine Mechanization and Automation, South African Institute of Mining and Metallurgy.
- [8] Tsai, R.Y., An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Miami Beach FL, 1986, pp. 364-374.
- [9] Zhang, Z., A flexible new technique for camera calibration. EEE transactions on Pattern Analysis and Machine Intelligence, 2000.
- [10] Johan C. van den Heuvel, J.C.M.K., Obstacle Detection For People Movers Using Vision And Radar. TNO Physics and Electronics Laboratory Oude Waalsdorperweg 63, The Netherlands, 2003.
- [11] Fusiello, E.T., and A. Verri, A compact algorithm for rectification of stereo pairs", Machine Vision and Applications. 12(1) 2000, pp.16-22.
- [12] Kuhl, A., Comparison of Stereo Matching Algorithms for Mobile Robots. The University of Western Australia Faculty of Engineering, Computing and Mathematics, 2005.
- [13] Sharkasi, A.T., Stereo Vision Based Aerial Mapping Using GPS and Inertial Sensors. Blacksburg, Virginia, 2008.
- [14] Pan, A.K.a.J., Purdue Experiments in Model-based Vision for Hallway Navigation. Proceedings of workshop on Vision for Robots in IROS'95 conference, Pittsburgh, PA, 1995, pp. 87-96.
- [15] Gonzalez, R.C., Digital Image Processing using Matlab. Pearson, Prentice Hall, 2002, pp 65-104

Rostam Affendi Hamzah received the B.Eng. in Universiti Teknologi Malaysia and currently pursues his M.Sc. in electronic system design engineering at Universiti Sains Malaysia,Malaysia. His current research interests include stereo vision, image processing and embedded system.

Sani Irwan Md Salim is a lecturer and received the B.Eng. in Universiti Teknologi Malaysia and M.Sc. in computer and communication engineering from QUT, Brisbane Australia. His current research interests embedded system computer engineering and real time application.