3D IMAGE PLANE FROM STEREO CAMERA CALIBRATION ON EXTRINSIC PARAMETERS IN STEREO VISION APPLICATION

R. A. Hamzah[1], N. M. Z. Hashim[2], H. N. Rosly[3], M. S. Hamid[4], K. A. A. Aziz[5]

Dept of Computer Engineering,
Faculty of Electronic and Computer Engineering
UTeM, 76100 Melaka, Malaysia

Email: [1]rostamaffendi,[2]nikzarifie,[3]hasrul,
[4]saad,[5]khairulazha@utem.edu.my

Abstract

This paper presents a 3D image plane in a group of target or image during the process of stereo pair calibration. The extrinsic parameters of camera calibration can be viewed in 3D image or scene which contains the rotation and translation of vector. The error re-projection of a single image could determine the less error of distortion during the extraction of chessboard corner each image taken. The distortion model also generates an error coordinate system in pixel value. The 3D image will viewed the result and output of extrinsic parameters during the calibration process.

Keywords: *camera calibration; Tsai's algorithm; rotation; stereo geometry; stereo camer; image translation*

I. INTRODUCTION

All the process of camera calibration is normally performed by first assuming a simplified model for both the camera and the distortion of resulting images and then statistically, usually fitting the distortion model to the observed distortion. Once the distortion has been modeled, it can be applied to the distorted image to correct it. An idealized pin hole camera model and radially symmetric lenses distortion are the usual modeling assumptions. Many calibration techniques also take into account distortion created by the digitization effects of the stereo camera. These effects are modeled along with the optical distortion and together account for the overall distortion observed in the acquired digital image. Once a model of the camera and distortion have been chosen they are fit to the real stereo camera by comparing where points of accurately know real world coordinates appear in the image and where they would appear if there was no distortion [1].

The error is minimized over as many points as is reasonable to fit distortion correction function to the distortion observed in the test scene. This distortion correction function, when applied to a raw image, reduces the distortion and what appears as a straight line in the real world appears as a straight line in the image [2]. In order for stereo correspondence techniques to work properly and for the range results that they yield to be precise and representative of the real world, the effects of the stereo camera distortion must repeatedly be accounted for. Usually, stereo vision systems use cameras that are horizontally aligned. That is, cameras are placed at the same elevation as shown in Figure 1.

II. CAMERA CALIBRATION

Typically, in stereo vision system the two cameras are very important. It means, these cameras should have the same characteristic. The selections will depend on the same size of pixel and manufacturer. After that, the installation of these two stereo cameras on the stereo application such as robot or arm robot, they have to be adjusted to get the aligned pictures or images. The installation of stereo pair in this paper is setup horizontally aligned with the range in between is six centimeters. That is, cameras are placed at the same elevation. The process of stereo vision is then usually defined as finding a match between features in left and right images as shown in Figure 2. In this paper, the horizontal baseline is used where the cameras are placed side by side of each other. In this case, the stereo vision consists of finding match between left and right image. In reality, the cameras will not have absolutely aligned optical axes. So, the images taken with this pair of stereo vision cameras will also contain some distortions.

Figure 1: Horizontally aligned of stereo camera.

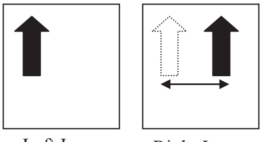


Figure 2: Features from left image are matched to features in the right image.

After stereo cameras installation, these cameras have to be calibrated. In order for stereo correlation techniques to work accurately and for the range results that they yield to be accurate and representative of the real world, the effects of the camera and lens distortion must often be accounted for. The process of camera calibration is generally performed by first assuming a simplified model for both the camera and the distortion of resulting images and then statistically, fitting the distortion model to the observed distortion. Once the distortion has been modeled, it can be applied to the distorted image to correct it [3]. The image processing software will use this result to get the accurate disparity values. Once a model of the camera and distortion have been chosen they are fit to the real camera and lens by comparing where points of accurately know real world coordinates appear in the image and where they would appear if there was no distortion. The error is minimized over as many points as is reasonable to fit distortion correction function to the distortion observed in the test scene [4]. This distortion correction function, when applied to a raw image, reduces the distortion and what appears as a straight line in the real world appears as a straight line in the image.

A. Method of Camera Calibration

Note that the purpose of camera calibration is trying to improve the transformations, based on measurements of coordinates, where one more often uses known transformation to map coordinates from one coordinate system to another. Tsai's method for camera calibration recovers the interior orientation, the exterior orientation, the power series coefficients for distortion, and an image scale factor that best fit the measured image coordinates corresponding to known target point coordinates. This is done in stages; starting off with closed form least squares approximation of some parameters and ending with an iterative non-linear optimization of all parameters simultaneously using these estimates as starting values [5].

Importantly, it is error in the image plane that is minimized. Details of the method are different for planar targets than for targets occupying some volume in space. Accurate planar targets are easier to make, but lead to some limitations in camera calibration [4]. The flowchart Figure 3 below shows the steps of stereo pair calibration's programming in matlab [6].

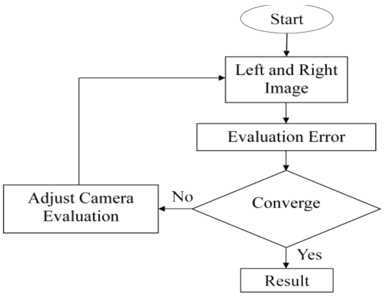


Figure 3: Flowchart of Tsai method for Extrinsic Parameters

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 [7]. This paper presents the calibration of stereo camera with the rotation stage. The calibration of rotation stage facilitates registration of multi-view range images into a common coordinate system. This paper using the Tsai's calibration technique using a calibration pattern which contains several controls points [8]. A focus calibration with respect to multiple distances of an object is also presented. A stereo pair is installed on the translation stage to implement the parallel stereo. Calibration of the stereo camera is done by Tsai's algorithm [8]. Consider a point P in Figure 4 in 3D space, its representations P_l and P_r in the left and the right camera coordinate systems. Then the transformation between two coordinate systems is:

$$P_r = R_l^r (P_l - t_l^r)$$

Where, R_l^r and t_l^r are rotation and translation matrices from the left to the right camera coordinate system, respectively. The translation stage moves the camera in horizontal direction x-axis in camera coordinate system [9]. Therefore, there is little error on epipolar line so that it can search stereo correspondence on epipolar line in horizontal direction. When there is an error in vertical direction y-axis in camera coordinate, it can be manually adjust the level of the translation stage to minimize it. Figure 4 shows the geometry of the stereo camera system. The baseline of the stereo camera B is 60 mm and the focal length of the camera f is 15.35 mm. The focal length was calibrated using Tsai's non-coplanar camera calibration technique. Consider again a 3D point P and its projection point on the right plane as shown in Figure 4.

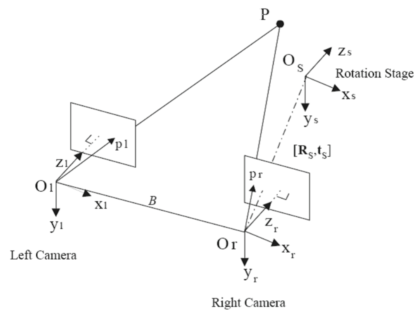


Figure 4: Stereo geometry of camera calibration

If finding another point pl in the left image plane which is the corresponding point of pr computed by stereo matching algorithm the vector *P* in 3D space can be conclude as and x_{pl} and x_{pr} are the x coordinates of the points pl and pr respectively.

$$P = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \dfrac{Bf}{x_{pl} - x_{pr}} \\ \dfrac{zx_{pr}}{f} \\ \dfrac{zy_{pr}}{f} \end{pmatrix}$$

B. Rotation Stage Calibration

Rotation stage calibration is important for registering multiple range images. This calibration gives transformation parameters between the camera coordinate system to the rotation coordinate system. These parameters are later used to register all partial 3D shapes into a common coordinate system and to integrate them. An accurate calibration of the rotation stage gives better performance for registration and integration of multiple range images. A calibration pattern which has several control points is used for calibrating the stage. Since several stereo images of an object are taken at every θ degree interval around an object, two consecutive images of calibration pattern are also taken with θ angle difference. By estimating calibration parameters which register two sets of 3D control points as close as possible, the partial shapes of the object into a common coordinate system also can be registered. Calibration for each view coordinate system is done using Tsai's algorithm. A checkerboard calibration pattern is placed on the rotation stage [9].

Two sets of stereo images are taken with θ degree angle difference. Let V_0 and V_1 denote the coordinate systems of the two views of directions and V_w denote the world coordinate system. Let a control point in Vk br P_k, where k=0 and 1, and Pw be the same point represented by the world coordinate system. Then transformations T_w^k from the world coordinate system to each camera coordinate systems are:

$$P_0 = T_w^0 P_w$$

$$P_1 = T_w^1 P_w$$

$$P_0 = T_w^0 (T_w^1)^{-1} P_1$$

$$= R_1^0 P_1 + t_1^0$$

Where R_1^0 and t_1^0 are rotation and translation matrices between two coordinate system as shown in Figure 5. In order to register all partial shapes to a common view coordinate system, it is necessary to find the transformation between the common view coordinate system view 0 (V0) in the rotation stage coordinate system. Let R_s and t_s be the rotation and the translation matrices between the rotation stages coordinate system and the camera coordinate system. The transformation between two view directions is also expressed as

$$P_0 = R_\theta R_S (P_1 - t_S) + t_s$$

$$P_0 = R_\theta R_S P_1 + t_s (I - R_\theta R_S)$$

From equation above,

$$R_s = R_\theta^{-1} R_1^0$$

$$t_s = (I - (R_1^0)^{-1}) t_1^0$$

C. Extrinsic Parameters

To obtain both extrinsic parameters from Figure 6 [10] of the stereo camera system, the Tsai method uses a chess board as calibration pattern. The process of camera calibration can be divided into three steps. Firstly the image acquisition for twenty images, then the extraction of the chess board corners in each image with 4x4 matrixes from Figure 7 and finally compute the external parameters value. In camera calibration process for stereo vision, the reference line of images is look at the epipolar line. A simplifying assumption for stereo vision is that epipolar geometry exists and that features can be matched along epipolar lines. For perfectly aligned cameras these lines are parallel to the baseline. The rotation and translation of vector is the output of extrinsic parameters.

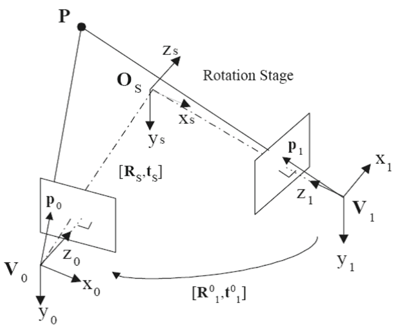


Figure 5: Multiview geometry of stereo coordination

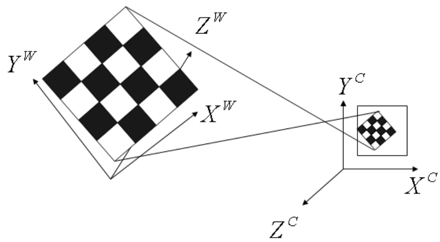


Figure 6. Rotation and translation of Extrinsic parameters

These output will be viewed in3D with the stereo camera location.

Figure 7: An extraction of chess board corner about 4x4

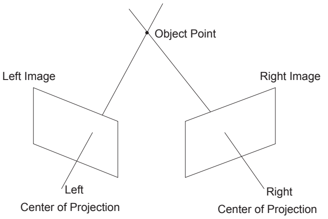


Figure 8: The projection of stereo image

III. 3D TRANSFORMATION MATRIX

The basic element of the stereo vision theory is triangulation [11](Wong, 1975). As shown in Figure 8, a 3D point can be reconstructed from its two projections by computing the intersection of the two space rays corresponding to it. The 3D location of that point is restricted to the straight line that passes through the center of projection and the projection of the object point. Binocular stereo vision determines the position of a point in space by finding the intersection of the two lines passing through the center of projection and the projection of the point in each image. In this section, this paper describes a calibration technique of a vision camera. It can be considered as an estimation of a projective transformation matrix from the world coordinate system to the camera's image coordinate system. For example the $(x_i^w y_i^w z_i^w, 1)$ of a 3D point in space and coordinates $(x_c, y_c, 1)$ of its projection on the 2D image plane, a 3X4 matrix M can be written according to the equation:

$$\begin{bmatrix} u_i \\ v_i \\ w_i \end{bmatrix} = M \begin{bmatrix} x_i^w \\ y_i^w \\ z_i^w \\ 1 \end{bmatrix}$$

With

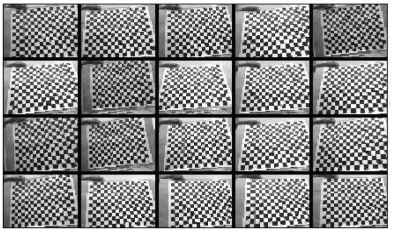
$$x_c = \frac{u_i}{w_i} = \frac{m_{11}x_i^w + m_{12}y_i^w m_{13}z_i^w + m_{14}}{m_{31}x_i^w + m_{32}y_i^w m_{33}z_i^w + m_{34}}$$

$$y_c = \frac{u_i}{w_i} = \frac{m_{21}x_i^w + m_{22}y_i^w m_{23}z_i^w + m_{24}}{m_{31}x_i^w + m_{32}y_i^w m_{33}z_i^w + m_{34}}$$

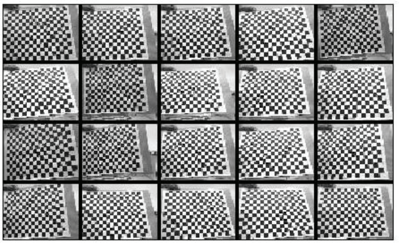
The matrix M is defined up to an arbitrary scale factor and has only 11 independent entries. Therefore it needs at least 6 world image points and their matching points in the image plane. If the calibration pattern is used, for example a checkerboard pattern, it have more correspondences and M can be estimated through least squares techniques.

IV. EXPERIMENT RESULT

From the experiment of stereo camera calibration, twenty images are captured from left camera and twenty images from right camera at the same time every capturing process. This is shown by Figure 9 the raw images in sequence. Figure 10 is the re-projection error mapping with the example of image 15 that randomly taken with its pixel (green color) coordinate (162.38, 87.43) and the pixel error is about (0.63324,-0.41004). Figure 11 is the 3D view of image plane with twenty chessboard position. It is split analysis of left and right camera. The third figure of Figure 11 is the whole 3D image with twenty different positions of targets. For Figure 12, first figure shows the impact of the complete distortion model (radial + tangential) on each pixel of the image. Each arrow represents the effective displacement of a pixel induced by the lens distortion. The second figure shows the impact of the tangential component of distortion. On this plot, the maximum induced displacement is 0.26, 023 pixel (at the upper left corner of the image). Finally, the third figure shows the impact of the radial component of distortion. On the three figures, the cross indicates the center of the image, and the circle the location of the principal point. Table 1 is the result of extrinsic parameters for camera calibration which contains of rotation and translation of images.



(a) Left Image



(b) Right Image

Figure 9:Raw image from stereo camera

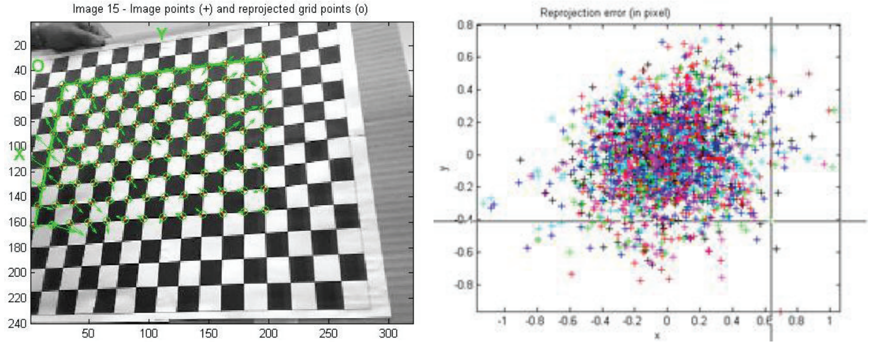


Figure 10: The example projection error of raw image

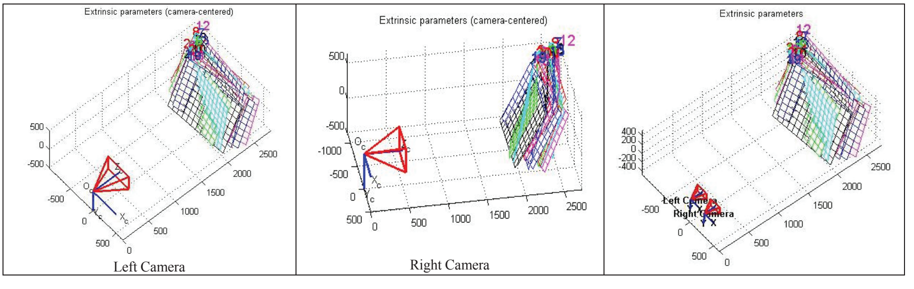


Figure 11: 3D image of stereo camera and chessboard position for extrinsic parameters

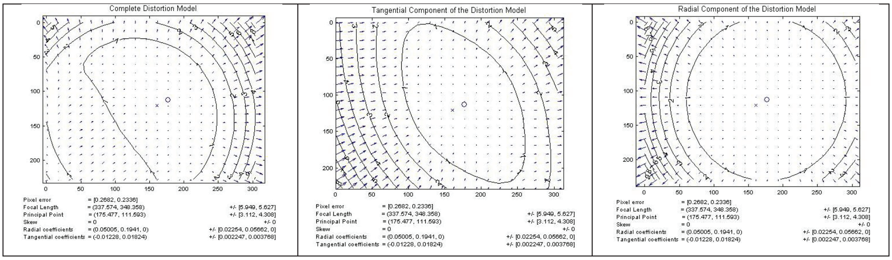


Figure 12: Visualization of image distortion during calibration process

TABLE I: EXTRINSIC PARAMETERS

Rotation vector (om):	[0.03752, 0.01455, 0.02346]
Translation vector (T):	[-303.95819, -45.55713, -74.97660]

V. DISCUSSION

With two cameras for extrinsic parameters, each camera gives a special 3D back projection line. These two back projection lines usually match up at exactly one point (rectification process). So, given a stereo setup it is possible to find the 3D position of a point by observing its position in two different cameras. It is possible that the point is infinitely far away and if the cameras are looking in the same direction separated only by a translation like binoculars, then the back projection lines are parallel, and will not (strictly speaking) intersect. However, if homogenous points are used carefully, an "infinite point" of the form will be recovered, which can be used to compute the direction of the point [11].

VI. CONCLUSION

The image plane works as a reference to the position of calibration target. This plane produces the rotation and translation vector for rectification of images in stereo vision analysis. The 3D image of target plane especially using the chess board will ensure the plane is well organized or in structured between the two cameras during a calibration process. The position for each other is ideally identical.

ACKNOWLEDGMENT

The authors would like to acknowledge the funding from Short Term Grant of Universiti Teknikal Malaysia Melaka.

REFERENCES

[1] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955.

[2] S. Rusinkiewicz, O. Hall-Holt, and M. Levoy, \Real-Time 3DModel Acquisition," Proceedings of SIGGRAPH 2002.

[3] CMU/FRC "Final Report: Autonomous Cross Country Navigation Technologies" Field Robotics Center or the Robotics Institute at Carnegie Mellon University, 1996.,

[4] 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.

[5] B. Curless and M. Levoy, \Better Optical Triangulation through Spacetime Analysis," IEEE International Conference on Computer Vision, pp. 987-994, June 1995.

[6] R A Hamzah, H. N Rosly, R. Rahim, Depth Evaluation in Selected Region of Disparity Mapping for Navigation

of Stereo Vision Mobile Robot, IEEE Symposium on Industrial Electronics & Applications, Penang, Malaysia, 2010,

[7] Rovira-Más, F., Q. Zhang, and J. F. Reid., 2004. Automated agricultural equipment navigation using stereo disparity images. Trans. ASAE 47(4).

[8] Tsai, R, "A Versatile Camera Calibration Technique for High Accuracy 3D Machine Vision Metrology Using Off the Shelf TV Cameras and Lenses", IBM Research Report, RC 11413, 1985.

[9] R. Y. Tsai, \A Versatile Camera Calibration technique for High-accuracy 3D Machine Vision Metrology using O_-the-shelf TV Camera and Lenses," IEEE Journal of Robotics and Automation, Vol. 3, No. 4, pp. 323-344, 1987.

[10] S. Thrun, G.B., Daniel Russakoff, Camera Calibration. Stanford CS223B Computer Vision, 2003.

[11] P. Hillman, Camera Calib and Stereo Vision, Edinburgh, UK, 2005