A 3-D VISION-BASED MEASUREMENT SYSTEM FOR ROBOT CALIBRATION

José Maurício S. T. Motta
Universidade de Brasília, Departamento de Engenharia Mecânica
Grupo de Automação e Controle - GRACO
70910-900 - Brasília, DF, Brazil.
Robert S. McMaster
Cranfield University, School of Industrial and Manufacturing Science
Building 62, Cranfield, MK43 OAL, England, UK.

Abstract. One of the problems that slow the development of off-line programming is the low static and dynamic positioning accuracy of robots. Robot calibration improves the positioning accuracy and can also be used as a diagnostic tool in robot production and maintenance. A large number of robot measurement systems are now available commercially. Yet, there is a dearth of systems that are portable, accurate and low cost. In this work it is presented a measurement system that can fill this gap in local calibration. The measurement system consists of a single CCD camera mounted on the robot tool flange with a wide angle lens, and uses space resection models to measure the end-effector pose relative to a world coordinate system, considering radial distortions. Scale factors and image center are obtained with innovative techniques, making use of a multiview approach. The target plate consists of a grid of white dots impressed on a black photographic paper, and mounted on the sides of a 90-degree angle plate. Results show that the achieved average accuracy varies from 0.2mm to 0.4mm, at distances from the target from 600mm to 1000mm respectively, with different camera orientations.

Keywords: Robot calibration, 3-D Machine Vision, Absolute accuracy, Camera calibration, Robotic Measurement Systems.

1. INTRODUCTION

Robot calibration plays an increasingly important role in robot production as well as in robot operation and integration within computer integrated manufacturing or assembly systems (Hidalgo & Brunn, 1998). The production, implementation and operation of robots are issues where robot calibration results can lead to significant accuracy improvement and/or cost-savings.

Theoretically, with the availability of robot calibration systems integrated with off-line programming systems, it would be possible to implement off-line programming in industry (currently they are not ready to be used), where multiple robots are used, by programming only one robot and copying the programs from one robot to the others. One of the most evident advantages of this type of programming is the reduction in maintenance expenditure, once the robot can return faster to service and easy reprogramming means the software becomes operational with little effort.

Robot calibration is an integrated process of modeling, measurement, numeric identification of actual physical characteristics of a robot, and implementation of a new model.

The second step of the calibration process is necessary for the determination of the "best fit" set of parameters to the actual robot. The types of measurement required comprise tool positions corresponding to joint positions. Since the fitting cannot be of better quality than that of the original data collection, it is very important to obtain a good set of data.

Cameras and Vision systems have become standard automation components. A typical robotic cell may feature an integrated multiple-camera system for product inspection, part orientation, identifying parts, searching for specific defects, and real-time monitoring of assembly accuracy (Asfahl, 1992).

Vision-based measurement systems using a single camera for robot calibration systems seem to be one of the trends of the development of robot calibration systems today, but the knowledge involved in this area is still restricted to research groups sponsored by large budget companies, with little interest to publish technical details (Albada *et al.*, 1995). For this technology to be available and used extensively in industry, there is still a long way to go.

This work presents the stages of the development of a feasible low cost vision-based measurement system using a single camera placed on a robot tool flange. Experimental results to assess the measurement system accuracy are presented as a step of the investigation of theoretical and practical aspects of the integration between the measurement system and the prototype of a robot calibration system.

2. PHOTOGRAMMETRIC MODEL

There are basically two typical setups for vision-based robot calibration. The first is to fix cameras in the robot surroundings so that the camera can frame a calibration target mounted on the robot end-effector. The other setup is named hand-mounted camera robot calibration. This latter setup can use a single camera or a pair of cameras. A single moving camera presents the advantages of a large field-of-view with a potential large depth-of-field, and a considerable reduced hardware and software complexity of the system. On the other hand, a single camera setup needs full camera re-calibration at each pose.

The goal of camera calibration is to develop a mathematical model of the transformation between world points and observed image points resulting from the image formation process. The parameters which affect this mapping can be divided into three categories (Prescott & McLean, 1997, Zhuang & Roth, 1996) : a) extrinsic (or external) parameters describe the relationship between the camera frame and the world frame, including position (3 parameters) and orientation (3 parameters); b) intrinsic (or internal) parameters describe the characteristics of the camera, and include the lens focal length, pixel scale factors, and location of the image center; c) distortion parameters describe the geometric nonlinearities of the camera. Some authors include distortion parameters in the group of intrinsic parameters (Tsai, 1987, Weng *et al.*, 1992). Distortion parameters can be present in a model or not.

The algorithm developed here to obtain the camera parameters (intrinsic, extrinsic and distortions) is a two-step method based on the Radial Alignment Constraint (RAC) Model (Tsai, 1987). It involves a closed-form solution for the external parameters and the effective focal length of the camera. Then, a second stage is used to estimate three parameters: the depth component in the translation vector, the effective focal length, and the radial distortion coefficient. The RAC model is recognized as a good compromise between accuracy and simplicity, which means short processing speed (Zhuang & Roth, 1996). Some few modifications were introduced to the original RAC algorithm, to be explained through the text.

2.1 RAC-Based Camera Model

In Figure 1 the world coordinate system is {xw, yw, zw}; the camera coordinate system is {x, y, z}. The origin of the camera coordinate system is centered at *Oc*, and the z-axis coincides with the optical axis. (X, Y) is the image coordinate system at *Oi* (intersection of the optical axis with the front image plane) and is measured in pixels. (u, v) are the analog coordinates of the object point in the image plane (usually measured in meters or millimeters). (X, Y) lies on a plane parallel to the *x* and *y* axes. *f* is the distance between the front image plane and the optical center.
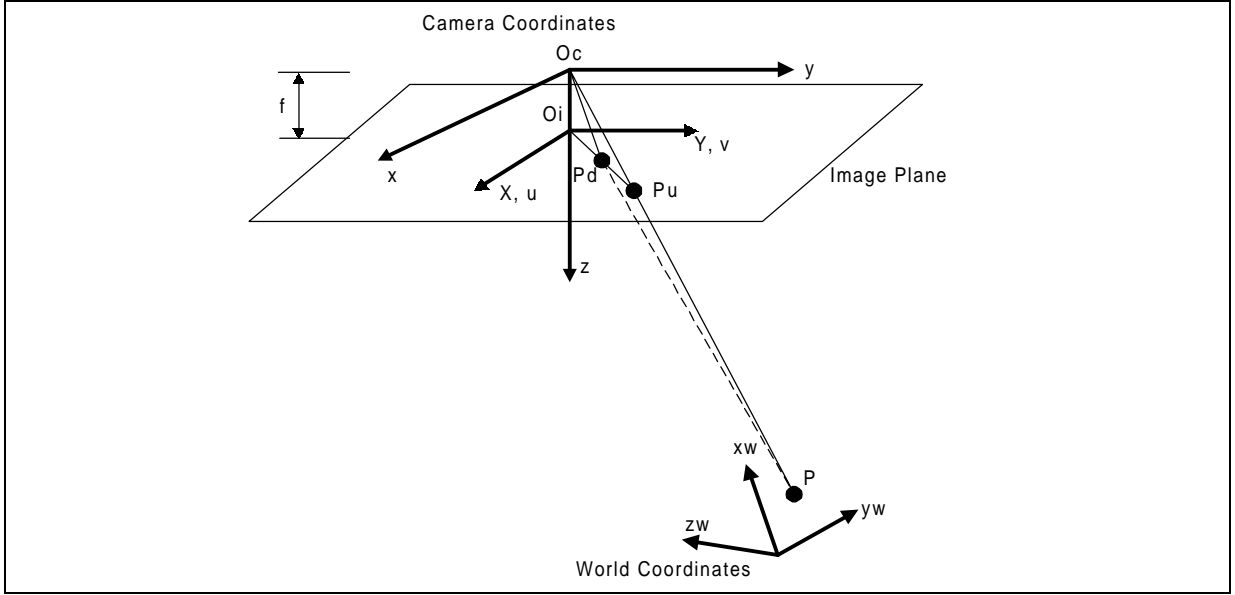

Figure 1 - Pin-hole model and the Radial Alignment Constraint hypothesis.

The rigid body transformation from the object coordinate system (xw, yw, zw) to the camera coordinate system (x, y, z) is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R. \begin{bmatrix} xw \\ yw \\ zw \end{bmatrix} + T \tag{1}$$

where R is the orthogonal rotation matrix, which aligns the camera coordinate system with the object coordinate system, and it can be represented as

$$R = \begin{bmatrix} r1 & r2 & r3 \\ r4 & r5 & r6 \\ r7 & r8 & r9 \end{bmatrix} \tag{2}$$

and T is the translation vector represented as

$$T = \begin{bmatrix} Tx \\ Ty \\ Tz \end{bmatrix} \tag{3}$$

which is the distance from the origin of the camera coordinate system to the origin of the object coordinate system represented in the camera coordinate system. (x_i, y_i, z_i) is the representation of any point (xw_i, yw_i, zw_i) in the camera coordinate system.

The distortion-free camera model, or "pinhole" model, assumes that every real object point is connected to its correspondent image on the image plane in a straight line passing through the focal point of the camera lens, *Oc*. In Figure 1 the undistorted image point of the object point, *Pu*, is shown.

The transformation from 3-D coordinate (x, y, z) to the coordinates of the object point in the image plane follows the perspective equations below (Wolf, 1983):

$$u = f . \frac{x}{z} \qquad \text{and} \qquad v = f . \frac{y}{z} \tag{4}$$

where *f* is the focal length (in physical units, e.g. millimeters).

The image coordinates (X,Y) are related to (u, v) by the following equations (Tsai, 1987, Zhuang & Roth, 1996):

$$X = s_u . u \qquad \text{and} \qquad Y = s_v . v \tag{5}$$

where s_u and s_v are scale factors accounting for TV scanning and timing effects and converting camera coordinates in millimeter or meters to image coordinates (X,Y) in pixels. Lenz and Tsai (1987) relate the hardware timing mismatch between image acquisition hardware and camera scanning hardware, or the imprecision of the timing of TV scanning only to the horizontal scale factor s_u. In Eq. (5), s_v can be considered as a conversion factor between different coordinate units.

The scale factors, s_u and s_v, and the focal length, *f*, are considered the intrinsic model parameters of the distortion-free camera model, and reveal the internal information about the camera components and about the interface of the camera to the vision system. The extrinsic parameters are the elements of R and T, which pass on the information about the camera position and orientation with respect to the world coordinate system.

Combining Eqs. (4) and (5) follows

$$X = s_u . u = f . s_u . \frac{x}{z} = fx. \frac{x}{z} \tag{6}$$

$$Y = s_v . v = f . s_v . \frac{y}{z} = fy. \frac{y}{z} \tag{7}$$

The equations above combined with equation (1) produce the distortion-free camera model

$$X = fx. \frac{r1. xw + r2. yw + r3. zw + Tx}{r7. xw + r8. yw + r9. zw + Tz} \tag{8}$$

$$Y = fy. \frac{r4. xw + r5. yw + r6. zw + Ty}{r7. xw + r8. yw + r9. zw + Tz} \tag{9}$$

which relates the world coordinate system (xw, yw, zw) to the image coordinate system (X,Y). *fx* and *fy* are non-dimensional constants defined in Eqs. (6) and (7).

Radial distortion can be included in the model as (Tsai, 1987, Weng *et al.*, 1992):

$$X.(1 + k.r^2) = fx. \frac{r1.xw + r2.yw + r3.zw + Tx}{r7.xw + r8.yw + r9.zw + Tz} \tag{10}$$

$$Y.(1 + k.r^2) = fy. \frac{r4.xw + r5.yw + r6.zw + Ty}{r7.xw + r8.yw + r9.zw + Tz} \tag{11}$$

where $r = \mu.X^2 + Y^2$ and μ is the ratio of scale to be defined further in the text. Whenever all distortion effects other than radial lens distortion are zero, a radial alignment constraint (RAC) equation is maintained. In Fig. 1, the distorted image point of the object point, Pd, is shown.

A mathematical "trick" to linearize the RAC model and find a linear least-squares solution to the camera calibration problem when radial distortion is considered was used by Zhuang & Roth (1993). Then, the RAC model with radial distortion expressed in equations (10) and (11) is changed to

$$\frac{X}{1 - k.r^2} \cong fx. \frac{r1.xw + r2.yw + r3.zw + Tx}{r7.xw + r8.yw + r9.zw + Tz} \tag{12}$$

$$\frac{Y}{1 - k.r^2} \cong fy. \frac{r1.xw + r2.yw + r3.zw + Ty}{r7.xw + r8.yw + r9.zw + Tz} \tag{13}$$

This transformation can be done under the assumption that $k.r^2 << 1$, i.e. $(1 + k.r^2) \cong (\frac{1}{1 - k.r^2})$ when $k.r^2 << 1$. The first stage to solve Eqs. (12) and (13) determines the rotation matrix R (Eq. 2), Ty and Tx. The algorithm for that is found in the Tsai's article (1987) and Zhuang & Roth's book (1996). The second stage is proposed here as

$$\begin{bmatrix} -X_i & x_i & -x_i.r_i^2 \\ -Y_i & \mu.y_i & -\mu.y_i.r_i^2 \end{bmatrix} . \begin{bmatrix} Tz \\ fx \\ k.fx \end{bmatrix} = \begin{bmatrix} X_i.w_i \\ Y_i.w_i \end{bmatrix} \tag{14}$$

where $x_i = r1.xw_i + r2.yw_i + Tx$, $y_i = r4.xw_i + r5.yw_i + Ty$, $w_i = r7.xw_i + r8.yw_i$, and zw_i is made null (all calibration points are coplanar). The overdetermined linear system in Eq. (14) can be solved by a linear least-square routine. The calibration points used to solve the system above were on the boarder of a square grid of 9x9 circular points with 1mm diameter, totting up 32 points. The grid has approximately 200x200mm.

2.2 Calibration of the Scale Factors

A ratio of scale is defined as

$$\mu = \frac{fy}{fx} = \frac{s_v}{s_u} \tag{14}$$

and dividing Eq. (12) by (13) yields

$$\frac{X}{Y} = \mu^{-1} . \frac{r1.xw + r2.yw + r3.zw + Tx}{r4.xw + r5.yw + r6.zw + Ty}$$
(15)

Two-step camera calibration techniques such as the RAC model can be used to determine the ratio of scale accurately if more than one plane of calibration points is used. This is accomplished by moving a single-plane calibration setup vertically with a z-stage (Zhuang & Roth, 1993). However, if only one coplanar set of calibration points is to be used, pre-calibration of the horizontal scale factor is essential.

Based on the fact that using non-coplanar and parallel planes μ can be determined from two-step calibration methods (Tsai, 1987), several images were obtained from different orientations and positions of the camera. This was accomplished when the camera was mounted on the robot hand during robot calibration measurements. Using the RAC model residuals as calculated from Eq. (15) (Lenz & Tsai, 1987):

$$r4.X_i.xw_i + r5.X_i.yw_i + X_i.Ty - r1.Y_i.xw_i - r2.Y_i.yw_i - Y_i.Tx = 0$$
(16)

where i is the index representing each calibration point in the image, and based on the fact that μ and the image center are quite independent each other in the RAC model (a poor guess for one does not affect the determination of the optimal value for the other) (Zhuang et $al.$, 1993), the residual found for each image were averaged and then μ was searched to minimize the average of the residuals. The solution found was $\mu = 0.9825$.

2.3 Calibration of the Image Center

The image center is defined as the frame buffer coordinates (Cx, Cy) of the intersection of the optical axis with the image plane. It is usually used as the origin of the imaging process and appears in the perspective equation. In high accuracy applications, it also serves as the center of radially modeled lens distortion (Zhuang & Roth, 1996). The image center is defined from

$$Cx = Xf - X \qquad and \qquad Cy = Yf - Y$$
(17)

where (Xf, Yf) is the computed image coordinates for an arbitrary point and (Cx, Cy) is the computed image coordinates for the center Oi in the image plane.

The Radial Alignment Constraint model holds true and is independent of radial lens distortion when the image center is chosen correctly. Otherwise, a residual exists and, unfortunately, the RAC is highly nonlinear in terms of the image center coordinates.

The method devised to find the image center was to search for the "best" image center as an average of all images in a sequence of robot measurements, using the RAC residuals. This method could actually find the optimal image center in the model, which could be easily checked calculating the overall robot errors after the calibration. The tests with a coordinate milling machine (explained further in the text) also showed that the determination of the image center by this method led to the best measurement accuracy in each sequence of constant camera orientation.

3. MEASUREMENT ACCURACY ASSESSMENT

The evaluation of the accuracy obtained by the vision measurement system was carried out using a Coordinate Milling Machine (CMM). The camera was fixed on the CMM's table and

moved on pre-defined paths as shown in Figure 2. The calibration board was fixed in front of the camera externally to the CMM, at a position that could allow angles from the camera optical axis to the normal of the calibration plane to be higher than 30° degrees (avoiding ill-conditioned solutions in the RAC model).

The CMM's table motion produced variations in z, x and y axes of the target plate relative to the camera coordinate system. Figure 3 shows the coordinate systems of the camera and calibration board.

There were three different measurement sequences of the 25 camera positions illustrated in Figure 2 by a cross mark. Each sequence was performed with the calibration board placed at different distances and orientations from the camera. The distances were calculated using the photogrammetric model.
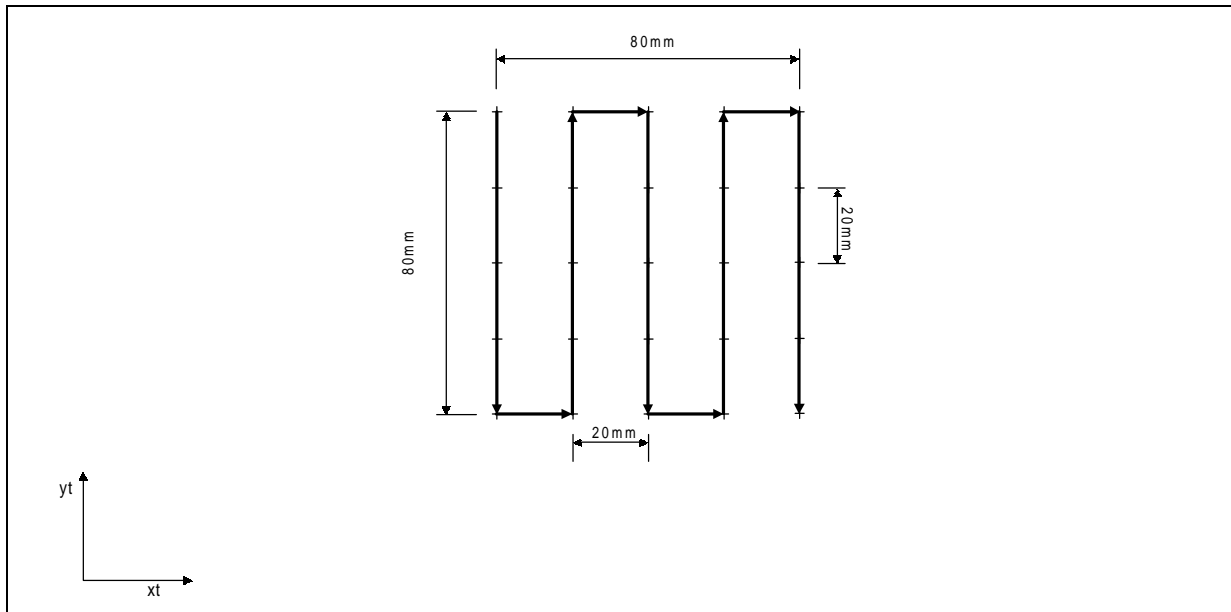

Figure 2 - Path of camera positions on the Coordinate Milling Machine table, during measurement accuracy evaluation of the vision system.

Once all external parameters of the camera (R, Tx, Ty and Tz) are determined for a particular camera pose in the camera coordinate system, it is necessary to establish an inverse transformation to determine the translation vector (XW, YW, ZW) from any point (on the calibration board) represented in the world coordinate system to the camera optic center. This translation vector has to be represented in the x, y and z directions of the world coordinate system. As defined in Eq. (1), R is the rotation matrix that aligns the world coordinate system with the camera coordinate system. From this, it is clear that

$$T = -R.\begin{bmatrix} XW + xw \\ YW + yw \\ ZW + zw \end{bmatrix} \qquad (18)$$

where T is the translation vector from the camera optic center to the origin of the world coordinate system (represented in the camera coordinate system) and (xw, yw, zw) is any point of the world coordinate system. This definition agrees with Eq. (1). Multiplying both sides of equation (18) by $-R^{-1}$ yields

$$\begin{bmatrix} XW \\ YW \\ ZW \end{bmatrix} = -R^{-1} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} \qquad\qquad (19)$$

where (x, y, z) is the representation of any point (xw, yw, zw) in the camera coordinate system, from Eq. (1). So, the vector (XW, YW, ZW) is the distance from any reference point on the calibration board to the optical center, represented in the world coordinate system.

The orientation of the line from the reference point to the optical center (target line), represented in the world coordinate system can be obtained from the rotation matrix
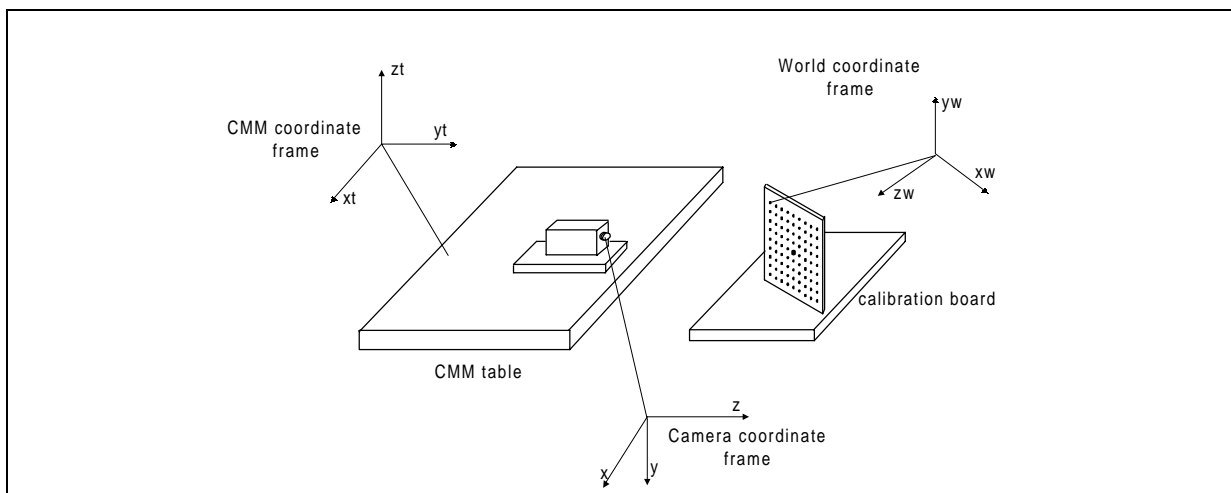
$$RW = -R^{-1} \qquad\qquad (20)$$

Figure 3 - Diagram showing coordinate systems and the experimental setup for the evaluation of the measurement system accuracy.

The calibration of the remaining intrinsic parameters of the camera, fx and k, was performed using the first sequence of 25 camera positions, which was chosen to be the closer to the target plate. For each image, values of fx and k calculated by the algorithm were recorded. Due to noise and geometric inaccuracies each image yielded different values for fx and k. The average values of the constants fx and k calculated from the 25 images were 1523 and 7.9×10^{-8} respectively. The standard deviation values for fx and k were 3.38 and 2.7×10^{-9} respectively.

The average values for fx and k should be kept constant from then on, and considered the "best" ones. To check this assumption, the distance travelled by the camera calculated by the Euclidean norm of (XW, YW, ZW) for each camera position, and the distance read from the CMM display were compared to each other. Errors were calculated to be the difference between the distance travelled by the camera using the two methods. It was observed then that the average value of fx did not minimize the errors.

Subsequently, an optimal value for fx was searched to minimize the errors explained above, by changing slightly the previous one. The value of k did not show an important influence on the errors when changed slightly. The optimal values of fx and k were checked in the same way for the other two measurement sequences at different distances, and showed to produce the best accuracy. The optimal values for fx and k where found to be 1566 and 7.9×10^{-8}, which were kept constant from then on.

Table 1 - Average, Median and Standard Deviation of the measurement system position accuracy calculated as the 3-D position error using the average rotation matrix.

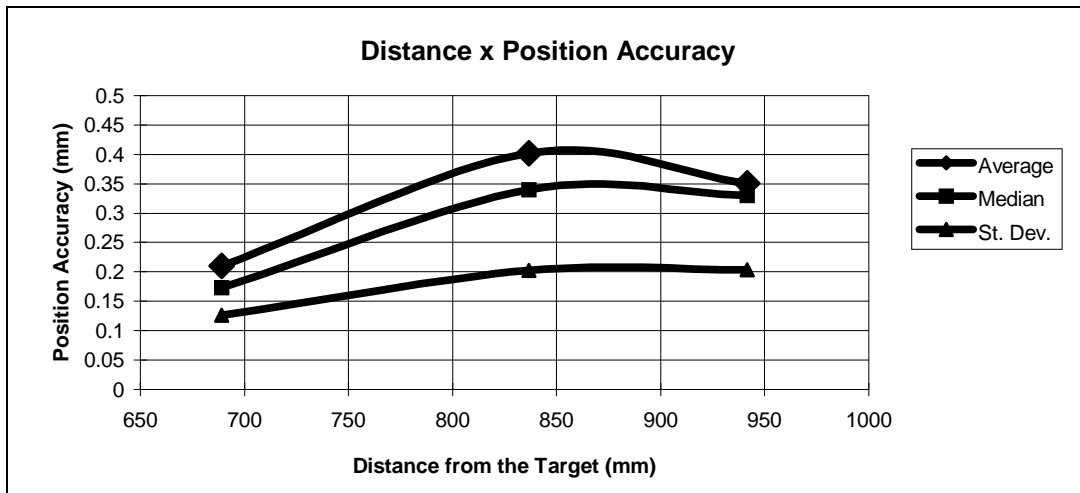| Measurement Sequence	Measurement Accuracy (mm)											
	Average				Median				Standard Deviation			
	x	y	z	3-D	x	y	z	3-D	x	y	z	3-D
1	0.13	0.11	0.07	**0.21**	0.10	0.10	0.09	**0.17**	0.10	0.09	0.06	**0.12**
2	0.18	0.30	0.12	**0.40**	0.13	0.26	0.11	**0.34**	0.12	0.20	0.09	**0.20**
3	0.17	0.20	0.16	**0.35**	0.10	0.18	0.13	**0.33**	0.17	0.11	0.13	**0.20**

Figure 4 - Measurement system position accuracy versus the average distance from the camera focal point to the central point of the target.

Assessment of 3-D measurement accuracy using a single camera is not straightforward. The method designed here to assess the measurement accuracy was to compare each component XW, YW and ZW corresponding to each camera position, to the same vector calculated assuming an average value for the rotation matrix. This method is justified considering that, since there are no rotations between the camera and the calibration board during the tests, the rotation matrix must be ideally the same for each camera position. Since the vector (XW, YW, ZW) depends on R in Eq. (2) and T in Eq. (3), and as T is calculated from R, all measurement errors would be accounted for in the rotation matrix, apart from the errors due to the ground accuracy of the calibration board. Of course, this assumption is valid only if another method to compare the camera measurements to an external measurement system (in physical units) is used to validate it. That means, the accuracy calculated from the travelled distance validates the accuracy calculated from the rotation matrix. The first method does not consider each error component independently (x, y, z), but serves as a basis to optimise the camera parameters. The second method takes into account the 3-D measurement error, assuming that the "correct" rotation matrix is the average of the ones calculated for each of the 25 positions. Table 1 shows the average, median and standard deviation of the 3-D position error for each component in *x*, *y* and *z*, calculated from the rotation matrix error.

The maximum position errors for each measurement sequence were 0.43, 0.97, and 0.72mm respectively. Figure 4 shows graphically the average, median and standard deviation values of the measurement system position accuracy calculated as explained before, as a

function of the average distance from the camera (optical center) to the central point of the calibration board.

4. CONCLUSIONS

Results showed that the RAC model used (with slight modifications) is not very robust, since even for images filling the entire screen and captured at approximately the same distances from the target, the focus length was not constant and showed an average value shifted by approximately 3% from the exact one. This amount of error can produce 3-D measurement errors much larger than acceptable. Practically speaking, the solution for this problem developed here for a set of camera and lens was to use an external measurement system to calibrate the camera, at least once. The measurement accuracy obtained is comparable to the best found in academic literature for this type of system, with median values of accuracy of approximately 1:3,000 when compared to the distance from the target. However, this accuracy was obtained at considerable larger distances and different camera orientations than usual applications for cameras require, making the system suitable for robotic metrology.

REFERENCES

Albada, G. D.; Visser, A.; Lagerberg, J. M. and Hertzberger, L. O., 1995, A Portable Measuring System for Robot Calibration, Proceedings of the 28th International Symposium on Automotive Technology and Automation - Dedicated Conference on Mechatronics and Efficient Computer Support for Engineering, Germany, 441-448.

Asfahl, C. R., 1992, Robots and Manufacturing Automation, John Wiley & Sons, Singapore, 2nd edition.

Hidalgo, F. and Brunn, P., 1998, Robot Metrology and Calibration Systems - A Market Review, Industrial Robot, Vol. 25, No. 1, 42-47.

Lenz, R. K. and Tsai, R. Y., 1987, Techniques for Calibration of the Scale Factor and Image Center for High Accuracy 3D Machine Vision Metrology, IN: Proceedings of the IEEE International Conference on Robotics and Automation, Raleigh, NC, 68-75.

Prescott, B. & McLean, G. F., 1997, Line-Based Correction of Radial Lens Distortion, Graphical Models and Image Processing, Vol. 59, No. 1, 39-47.

Tsai, R Y., 1987, A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the Shelf TV Cameras and Lenses, IEEE International Journal of Robotics and Automation, Vol. RA-3, No. 4, 323-344.

Weng, J.; Cohen, P. and Herniou, M., 1992, Camera Calibration with Distortion Models and Accuracy Evaluation, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 14, No. 10, 965-980.

Wolf, P. R., 1983, Elements of Photogrammetry, 2nd Edition, McGraw-Hill, Singapore.

Zhuang, H. and Roth, Z. S., 1993, A Linear Solution to the Kinematic Parameter Identification of Robot Manipulators, IEEE Transactions on Robotics and Automation, Vol. 9, No. 2, 174-185.

Zhuang, H. and Roth, Z. S., 1996, Camera-Aided Robot Calibration, CRC Press, USA.

Zhuang, H.; Roth, Z. S; Xu, X. and Wang, K., 1993, Camera Calibration Issues in Robot Calibration with Eye-on-Hand Configuration, Robotics & Computer Integrated Manufacturing, Vol. 10, No. 6, 401-412.