TIMS XXX-Sograpo XXXII, Joint International Meeting, Rio de Janeiro, 1991.

Calibration of a Tridimensional Vision System for Robotics

DIAS, Jorge BATISTA, Jorge ARAÚJO, Helder ALMEIDA, Traça de NUNES, Urbano

Departamento de Engenharia Electrotécnica
Universidade de Coimbra
Portugal

Abstract: *This paper addresses the geometric modelling aspects related with the use of a pair of video cameras mounted on a six degrees-of-freedom manipulator in an hand-eye configuration. The information obtained by these sensors shall be used to generate the depth map of the environment where the manipulator works. The procedure to determining the image formation parameters is known by camera calibration and a method to optimize these calibration parameters is presented in the paper. After present the calibration technique used, we develop a calibration matrix invariant to movement.The optimization process is applied to this matrix to obtain a better performance in the numerical data. This optimization process is the base for a recalibration procedure that exploits the capability of positioning of the cameras in different positions within the workspace of the* ·*bot. The first optimization process used is a recursive* ·*m of the least-squares method and the second is the* *Kalman filter. Experimental results obtained, shows that the recalibration technique gives some improvement in the numerical data and stabilizes the calibration matrix along the different positions of the vision system within the robot workspace.*

1. Introduction

To use a vision system in a robotic system is necessary to know the transformation between the camera space - 2D coordinates of the image - and the robot space - 3D coordinates of the world. If the robotic system is a manipulator and the vision system is attached to the last link, this transformation is normally known as hand-eye transformation. For the use of vision systems in robotics is fundamental to know this transformation. The evaluation of it implies to solve the problem of camera calibration.

The problem of camera calibration has been studied since the beginning of the research on computer vision and is ·e process of determining the parameters of the model of ·ge formation. The solution for this problem passes for establish the relationship between 3D world coordinates and their corresponding 2D image coordinates obtained by sampling the video signal the TV cameras used by visual system. This relationship is normally expressed by equations that relates the image coordinates (x_f, y_f) with the three-dimensional world coordinates (x_w, y_w, z_w). Using the pin-hole model for image formation the calibration process can be divided in two stages: first, the calibration of the intrinsic parameters and second, the calibration of the extrinsic parameters of the camera. The intrinsic parameters describe the geometry of the image formation and the extrinsic parameters describes the 3D position and orientation of the camera frame relative to a world coordinate system. Once these relationships are known, the three-dimensional information can be inferred from two-dimensional information using computer vision techniques like stereo, motion or focus.

Normally the calibration algorithms proposed in the literature are for cameras in fixed positions. These calibration methods can classified into five categories [12]: (1) techniques that involves non-linear optimization; (2) techniques that uses perspective model for image formation but use linear approximation for equation solving; (3) the method of two planes; (4) techniques based in some geometrical characteristics; (5) the method proposed by Tsai. Several new calibration techniques are recently referred [13,14,8] and proposed [5]. In [6] and [5], is suggested some improvement on the fixed position calibration techniques for eye-on-hand configurations, by optimization with Kalman filtering. An algorithm to calibrate dynamically an eye-on-hand system is presented in [9]. In this paper is developed a movement invariant transformation and also a method to improve the numerical results of this transformation by recursive least-squares or Kalman filtering and based on the method proposed by Tsai.

In the next section the image formation process is revised and the equations that describes and models the physical process is established. Section three, explains how that the parameters of the model are determined and in section four, the invariant form of camera calibration is developed. Using this invariant form the recalibation process is explained and, in section five, some experimental results are presented and analyzed.

2. Camera Model

The main problem in camera calibration is to find out, with a reasonable accuracy, the relative transformation between the 2D images and the referential of the three-dimensional world. In our case one pair of cameras in a stereo configuration is placed on the last link of the manipulator with a referential associated to it and known as {TOOL}. Two references called {CAM_LEFT} and {CAM_RIGHT} are placed in each camera and two more, {BASE} and {W} are created to represent the referential of the manipulator and the referential of the world respectively.

We begin by analyzing the left and right images of a grid of points whose coordinates in {W} are known. By using the correspondence between the points in the 3D scene and the corresponding points in the 2D image we can write a set of equations that models the image formation phenomena. The model used is based in perspective projection and assumes that the lens distortion is radial. This approach follows the same principle presented by Tsai [12] for the image formation, and uses the same relations between the different stages of the image formation.

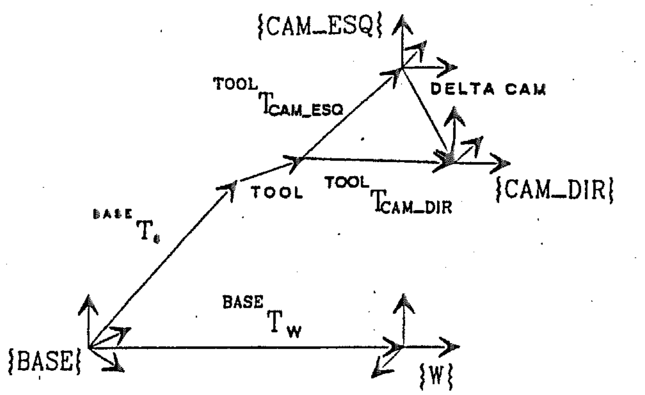


Figure 1 - Geometrical relations between the different referentials used in the system

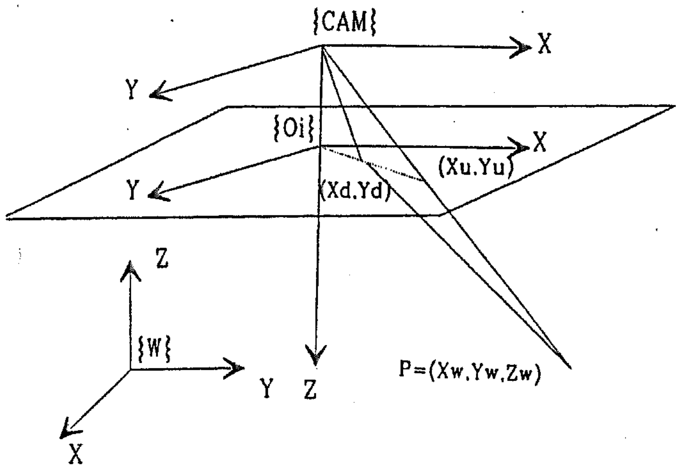


Figure 2 - The image formation model

Since the model used for each camera is equal, the study for the left and right cameras is equal and they can be represented by one common referential named {CAM}. Taking a point p in the three-dimensional space, that point it is represented in the referential {W} by (x_w, y_w, z_w) and by (x, y, z) in referential {CAM}. The origin of the referential {CAM} is coincident with the center of the lens. The image plane is parallel with the plane-xy of the {CAM} referential with an f coordinate corresponding to the focal length along the axis-z. In that image plane, a point p of the real world is represented by its projection (x_u, y_u) with the lens distortion neglected. Taking the distortion created by the lens as being radial, the projection changes from (x_u, y_u) to coordinates (x_d, y_d). In this way, the transformation between the referential {W} and the 3D referential {CAM} is given by:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \text{Rot} \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + \text{Trans} \qquad (2.1)$$

where (x, y, z) and (x_w, y_w, z_w) are the coordinates of the point p expressed in {CAM} and {W} respectively. The matrix **Rot** is one matrix (3x3) representing the rotation between the references {CAM} and {W} and **Trans** the (3x1) vector representing the translation between the same references. The relation between the points in the world and its own projections in the ideal image are given by the perspective transformation:

$$\begin{cases} x_u = f\frac{x}{z} \\ y_u = f\frac{y}{z} \end{cases} \qquad (2.2)$$

The relation between points in the ideal image to the related points in the real image is given by

$$\begin{cases} x_u = x_d \frac{D}{2} \\ y_u = y_d \frac{D}{2} \end{cases} \qquad (2.3)$$

where D is the radial distortion factor that depends of the distortion coefficient k_1 ([12] and [8]):

$$D = \frac{2}{1+k_1 R_d^2} \qquad \text{with } R_d^2 = x_d^2 + y_d^2$$

The relation of the real-image coordinates (x_u, y_u) to computer-image coordinates (x_f, y_f) is given by:

$$\begin{cases} x_f = S_x x_d + c_x \\ y_f = S_y y_d + c_y \end{cases} \qquad (2.4)$$

where
- S_x, S_y - are the scale factors for the x and y axes, respectively.
- (c_x, c_y) - are the computer-image coordinate for the origin in the image

The parameters $(S_x, S_y, c_x, c_y, f, k_1)$ are called the <u>intrinsic parameters</u> and **Rot** and **Trans** the <u>extrinsic parameters</u> that we need to evaluate for the left and right cameras.

3. Determination of the model parameters

The intrinsic and extrinsic parameters can be evaluated by the process presented in [12] using the model for image formation described. In that process the determination of the parameters is only for a camera in static position and is based in two stages: the first, to compute the extrinsic parameters corresponding to the 3D orientation **Rot** and the x and y components of **Trans**; the second, to compute the intrinsic parameters (f, k_1)and the z component of **Trans**.

The process begins by grabbing an image into the computer frame memory and using a scene with a set of points where their three-dimensional coordinates are known. After binarize the image, the correspondence between points in the image and their three-dimensional coordinates is done. This set of correspondent points is the input for our calibration algorithm where the parameters evaluation of the intrinsic and extrinsic parameters are evaluated.

4. Invariant form and recalibration

If the right and left cameras are calibrated in a predefined position using the process above and after the manipulator tool is moved the extrinsic parameters **Rot** and **Trans** are not more valid. Normally the intrinsic parameters should not change if we do not use auto-focus or controlled zoom. In the practice some small differences in the values are observed mainly due to non-linearities of the lens and the different illumination condition at different points of view.

The extrinsic parameters obtained by the calibration process are represented by the transformation ${}^{W}T_{CAM}$ which give the relation between the references of the right and left images and the 3D referential {W}. By another way, is possible to know with relative accuracy, the tool coordinates ${}^{W}T_{TOOL}$ using the programming facilities of the manipulator. Using this transformation we can determine ${}^{TOOL}T_{CAM}$ using the equation:

$$ {}^{TOOL}T_{CAM} = ({}^{W}T_{TOOL})^{-1}\,{}^{W}T_{CAM} \tag{4.1} $$

This transformation have the propriety of to be invariant to manipulator's tool movement. Using this propriety we can implement an algorithm to optimize the numerical values of the extrinsic and intrinsic parameters obtained by the calibration process for a static position, extended it to dynamical use.

Manipulating the equations (2.1), (2.2), (2.3) and (2.4) we can obtain explicit relations between extrinsic and intrinsic parameters for the cameras using the relation between 2D points in the computer image and the 3D points related to the system {CAM}, given by

$$ \begin{cases} \frac{x_f}{fS_x}z - \frac{c_x}{fS_x}z - \frac{x\,k_1}{S_x^2}(x_f^2 + c_x^2 - 2\,x_f c_x) - \frac{x\,k_1}{S_y^2}(y_f^2 + c_x^2 - 2\,y_f c_x) = x \\ \frac{y_f}{fS_y}z - \frac{cf_y}{S_y}z - \frac{x\,k_1}{S_x^2}(x_f^2 + c_x^2 - 2\,x_f c_x) - \frac{x\,k_1}{S_y^2}(y_f^2 + c_x^2 - 2\,y_f c_x) = y \end{cases} \tag{4.2} $$

and for the relation between 3D points expressed in referential {CAM} and the same points expressed in referential {TOOL},

$$ \begin{cases} f\frac{x}{z} = f\frac{\mathbf{p}^T\mathbf{r}_1 + T_x}{\mathbf{p}^T\mathbf{r}_3 + T_z} \\ f\frac{y}{z} = f\frac{\mathbf{p}^T\mathbf{r}_2 + T_y}{\mathbf{p}^T\mathbf{r}_3 + T_z} \end{cases} \qquad (4.3 $$

The vector (T_x, T_y, T_z) represents the origin coordinates of the referential {CAM} expressed in referential {TOOL} and $\mathbf{r}_i$ is the line i of the Rot matrix which describes the orientation of {CAM} relatively to the manipulator's tool.

After manipulating the equations (4.2) and (4.3) relating the image and three-dimensional points, it is straightforward to obtain an expression from which we can apply recursive methods for the estimation of the cameras' parameters using the information acquired from different positions of the stereo system. This procedure gives us a dynamic form to optimize (or filter) the parameters of the cameras. The mathematical tools used to implement such estimation are based on the covariance form of the recursive least-squares method and the Kalman filtering [10].

Beginning by the equation (4.3) and manipulating it algebraically we obtain the matrix form:

$$ \mathbf{A}\,\mathbf{m} = \mathbf{b} \tag{4.4} $$

where $\mathbf{m}$ is the vector of unknowns. For example the relation between the elements of $\mathbf{m}$ and the intrinsic parameters for the center of the image are given by:

$$ c_x = \frac{m_2}{m_0};\; c_y = \frac{m_3}{m_1} $$

Doing the same for equation (4.3), we obtain two matrix equations

$$ \begin{bmatrix} p^T & 0_{(1x3)} & \alpha p^T \\ 0_{(1x3)} & p^T & \beta p^T \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} = \begin{bmatrix} \alpha T_z - T_x \\ \beta T_z - T_y \end{bmatrix} \qquad \Rightarrow \mathbf{B}\,\mathbf{n} = \mathbf{c} \tag{4.5} $$

and

$$ \begin{bmatrix} -1 & 0 & \alpha \\ 0 & -1 & \beta \end{bmatrix} \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} = \begin{bmatrix} p^T r_1 - \alpha p^T r_3 \\ p^T r_2 - \beta p^T r_3 \end{bmatrix} \qquad \Rightarrow \mathbf{D}\,\mathbf{t} = \mathbf{e} \tag{4.6} $$

with $\mathbf{p}=(x_w, y_w, z_w)$, $\alpha=\frac{x}{z}$, $\beta=\frac{y}{z}$, $\mathbf{r}_i$ the line i of the Rot matrix, and (T_x, T_y, T_z) the elements of the Trans vector. The vectors $\mathbf{n}$ and $\mathbf{t}$ are the unknowns representing the extrinsic parameters.

Using the equations (4.4),(4.5) and (4.6) is easy to apply a recursive method of estimation the parameters. For each iteration of the recalibration recursive process, two stages are used: first, the recalibration of the intrinsic parameters using (4.4), and second, the recalibration of the extrinsic parameters using (4.5) and (4.6).

Recursive Least-Squares

According to [10], if a system can be represented by a linear equation:

$$ \mathbf{Z}_k = \mathbf{H}_k\,\mathbf{X} \tag{4.7} $$

where $\mathbf{H}$ and $\mathbf{Z}$ are matrices or vectors of the k measures read respectively in the input and output signals of the system, and $\mathbf{X}$ is the vector of the system unknown parameters, a recursive solution to obtain $\mathbf{X}$ by least-squares can be expressed by:

$$ \hat{\mathbf{X}}_{k+1} = \hat{\mathbf{X}}_k + \mathbf{Kw}_{k+1}[\,\mathbf{z}_k - (\mathbf{H}_k)^T\hat{\mathbf{X}}_k] \tag{4.8} $$

where $\mathbf{Z}_{k+1}$ and $\mathbf{H}_{k+1}$ are matrices obtained using the last $(k+1)$ measure.

The $\mathbf{Kw}_{k+1}$ coefficient is obtained by

$$ \mathbf{Kw}_{k+1} = \mathbf{P}_k\,\mathbf{H}^T_{k+1}[\,(\mathbf{H}_{k+1})^T\,\mathbf{P}_k\,\mathbf{H}_{k+1} + 1\,]^{-1} \tag{4.9} $$

which can be understood as a gain for the $(k+1)$ iteration and where the factor $\mathbf{P}_k$ is given by

$$P_k = [I - Kw_k (H_k)^{-1}]\, P_{k-1} \qquad (4.10)$$

The application of this method to the equations (4.4), (4.5) and (4.6) is direct. The initial conditions for $\hat{X}_0$ are given by a first calibration made for one position and the initial values for P_0 are given by:

$$P_0^{-1} = H_0 (H_0)^T$$

where $k=0$ signifies that no recalibrations were made.

<u>Kalman Filter</u>

The Kalman filter is an optimal procedure to optimize measures made in a continuously changing environment as the calibration of dynamic vision system.

Assume that X_{k+1} is the signal state to be estimated at time k+1 and our system can be expressed by the state matrices

$$Z_k = H_k X_k + v_k \qquad v_k \text{ is } N(0,R_k)$$
$$X_k = \Phi_{k-1} X_{k-1} + \omega_{k-1} \qquad \omega_k \text{ is } N(0,Q_k)$$

where v_k and ω_k are uncorrelated Gaussian white noise sequences with null mean and covariance matrices R_k and Q_k respectively.

The recursive equations of the Kalman filter which give a new estimate (X_k, R_k) of X and the covariance of the error based on the preceding values (X_{k-1}, R_{k-1}) are the following

$$X_k = X_{k-1} + Kw_k (Z_k - H_k X_{k-1})$$
$$Kw_k = P_{k-1} H_k^T [H_k P_{k-1} H_k^T + R_k]^{-1} \qquad (4.11)$$
$$P_k = (I - Kw_k H_k) P_{k-1}$$

These equations and the equations presented above for least-squares have same structure with a difference on the evaluation of the gain which ponders continuously the noise on the data using the covariance matrix. This recursive process is initialized by X_0 and by R_0 which are the "a priori" estimate of X and the covariance matrix associated with the error in the estimate.

Adapting these methods to the special conditions of (4.4), (4.5) and (4.6), the calculations for both methods, have the following development:

$$\cdots P_k \Rightarrow Kw_{k+1} \Rightarrow \hat{X}_{k+1} \Rightarrow P_{k+1} \cdots$$

5. Experimental Results

The experimental setup used was based in a integrated and evolutive system in which the addition of new sensors and actuators is possible in a modular fashion. The control system exhibits an hierarchical structure following a distributed processing model [1]. To establish the communications between the Supervisor Computer and the manipulator controller, a package of software implementing the DDCMP protocol was made [3]. The vision system is fastened to the last link of the manipulator by a special tool designed for that purpose and uses two CCD cameras. The two video signals are sampled and stored in memory using an AT-bus frame grabber and the images are sampled by 512x512 pixels. The software for the calibration process was developed in C and runs on the Supervisor Computer. All additional image processing necessary in the process use routines from a software library developed in our laboratory [2].

To evaluate the effect of the recalibration process we begin by taking different images of the calibration grid using different points of view. The three-dimensional points are then transformed to the referential {TOOL} and the correspondence between the images points and the three-dimensional world is established.

This set is the input data for both recalibration processes - recursive least-squares and Kalman filter - as we can see in figures below. Using as first approximation the result obtained by Tsai method for a static position, the difference between the application of the two processes it is not significant after the 10th point. The significant difference between the two methods is the speed of convergence, which it is bigger for Kalman filter. Nevertheless, this implies more unstable results as show for C_x and C_y diagrams.

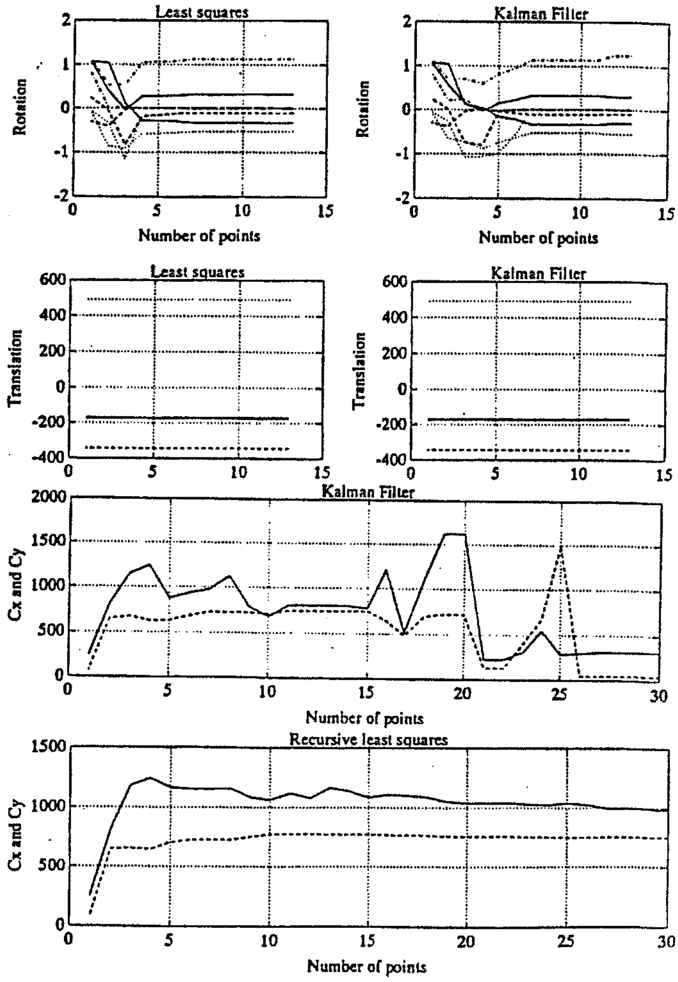


Figure 3 - The figure shows the effect of the use of more points in recalibration process

The estimation process involves the calculation of the dynamic gain Kw and the figure 4 exemplifies the evolution of it with the number of points used.

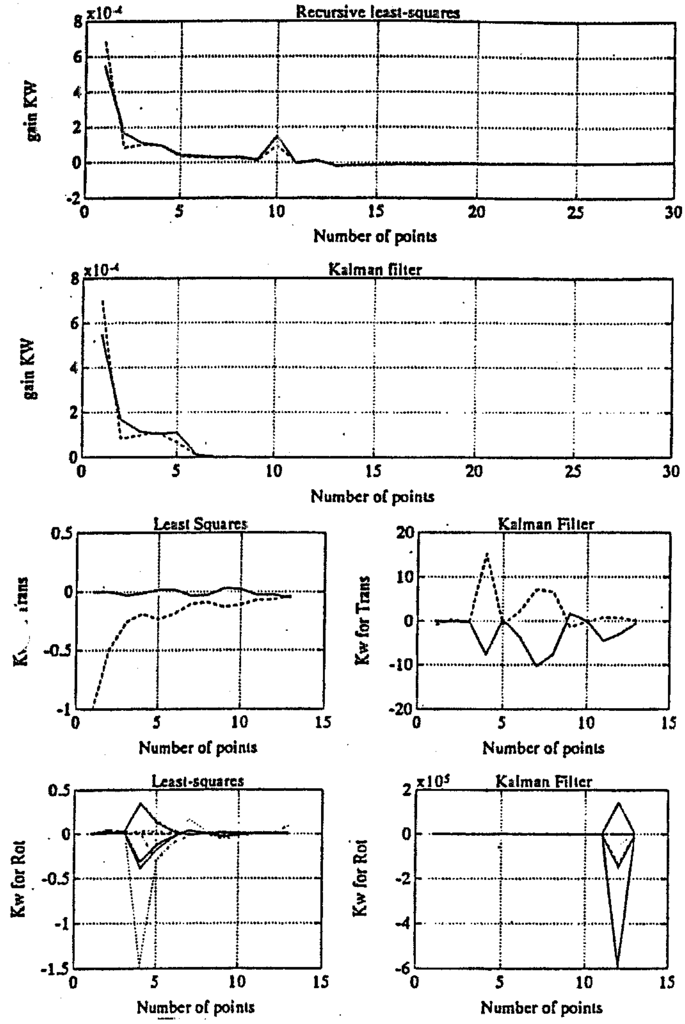
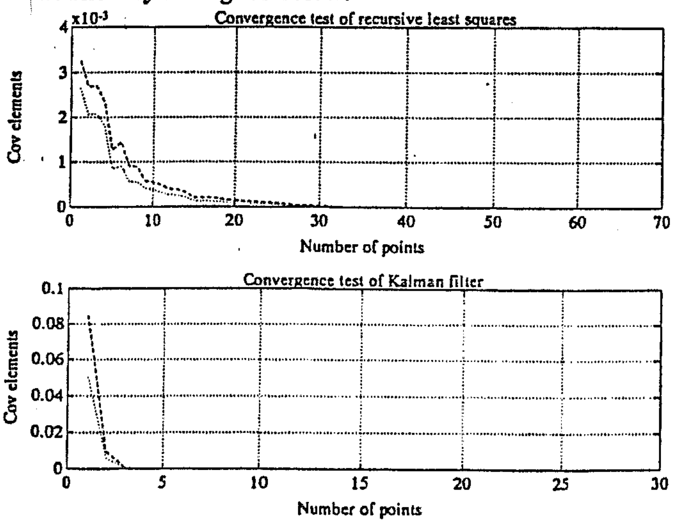


Figure 4- The diagrams represents the evolution of Kw with the number of points used. We can see that after some quantity of points the solution stabilizes for both the methods

When we use this recursive methods it is important to know if the process converges. One method is to trace the elements of the error-covariance matrix **P** associated with [illegible] as show by the figure below.



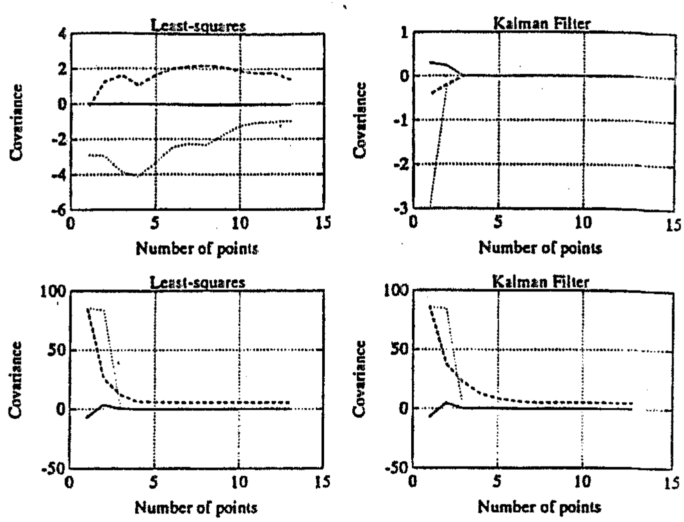


Figure 5 - The diagram shows the convergence of the two methods

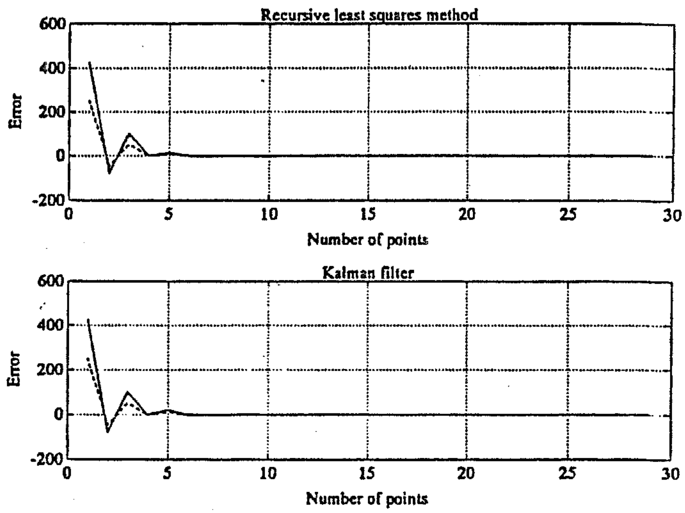


Figure 6 - The diagram shows that the difference between the real points and the backprojected points (the error) tends to zero after the 5th point and for both the methods

6. Conclusions

In this paper, the calibration of a stereo pair of cameras attached to the end-effector of a six degree-of-freedom manipulator was addressed. In particular a recalibration process based on a recursive least-squares estimation was proposed. This work represents an initial stage of a wider robotics research project having in view the integration of different visual techniques for building local and global descriptions of the 3D world and evaluation of visual-based path control strategies.

The method used for the calibration [12] is based on a priori knowledge of two intrinsic parameters which can be obtained by the distance between sensors in y and x directions of the sensor plate on CCD cameras. Normally these values are supplied by the manufacturer but the experience teach us that the values are not accurate enough. The experiments shows that the algorithm proposed by Tsai is very sensitive to the numerical value of these parameters. This problem is the objection to this approach (see also [11]) and the results obtained are not better that obtained by another methods [4,7]. This handicap can be overtake by the recalibration process.

The experiences made, shows that both the methods - recursive least-squares and Kalman filtering - are valid to use with recalibration. For better performance the choice should be the Kalman filter. As general result, the recalibration process give a very stabilized solution for the calibration matrix within the manipulator's workspace.

References

[1] A.T. Almeida, U.C. Nunes, J. Dias,; H. Araújo, J. Batista,: A Distributed System for Robotic Multi-Sensor Integration. Int.Journal of Industrial Metrology ,1990, pag. 217-229

[2] J. Batista, J. Dias, H. Araújo,.: Biblioteca de Programas para Processamento de Imagem. In: 3 Simpósio de Electrónica das Telecomunicações, Porto, 1988 (in portuguese)

[3] E. Capote; J. Machado, F. Bastos, J. Dias.: Utilização do DDCMP". Int.Technical Report, DEE-University of Coimbra, 1989 (in portuguese)

[4] Faugeras, O. D., Toscani, G.: Camera Calibration for 3D Computer Vision, In Proc of Int. Workshop on Machine Vision and Machine Intelligence, Tokyo, Japan, Feb. 1987

[5] W.I. Grosky and L.A. Tamburino, : A Unified Approach to the Linear Camera Calibration Problem. IEEE Trans. on Pat. Anal. and Mach. Intell., Vol.12, No.7, July 1990

[6] Y-P. Hung,: A Simple Real-Time Method for Calibrating a Camera Mounted on a Robot for Three Dimensional Machine Vision. SPIE Vol. 1005 Optics, Illumination, and Image Sensing for Machine Vision III, 1989

[7] A. Izaguirre, P. Pu, J. Summers: A New Development in Camera Calibration: Calibrating a Pair of Mobile Cameras. Proc. Int. Conf. on Robotics and Automation, 1985

[8] R.K. Lenz, and R.Y. Tsai,: Calibrating a Cartesian Robot with Eye-on-Hand Configuration Independent of Eye-to-Hand Relationship. IEEE PAMI, Vol .11, No 9, Sept. 1989

[9] P. Liang, Y.L. Chang, S. Hackwood: Adaptive Self-Calibration of Vision-Based Robot Systems. IEEE Trans. on Sys. Man and Cyb., Vol.19, Nº4, July/Aug. 1989

[10] J.M. Mendel,: Lessons in Digital Estimation Theory. Prentice-Hall Signal Processing Series, New-Jersey, 1987, ISBN 0-13-53 0809-7 025

[11] P. Puget, T. Skordas, An Optimal Solution for Mobile Camera Calibration, in First European Conference on Computer Vision, April 1990

[12] R.Y. Tsai: A Versatile Camera Calibration Technique for High Accuracy 3D Machine Vision Metrology Using off-the-shelf TV Cameras and Lenses. IBM RC 11413, October 1985

[13] R.Y. Tsai, and R.K. Lenz,: Overview of a Unified Calibration Trio for Robot Eye, Eye-to-Hand, and Hand Calibration Using 3D Machine Vision. Proc. SPIE-Sensor Fusion: Spatial Reasoning and Scene Interpretation, 7-9 Nov. 1988

[14] R.Y. Tsai, and R.K. Lenz,: A New Technique for Fully Autonomous and Efficient 3D Robotics Hand/Eye Calibration. IEEE J. of Robotics and Automation, Vol. 5, No. 3, June 1989