

A FILTERING-BASED APPROACH TO EYE-IN-HAND ROBOT VISION

Antonio Maria Garcia Tommaselli
DCart-FCT-Unesp - Pres. Prudente
e.mail: uepr@brfapesp.bitnet
Brazil

Clésio Luís Tozzi
DCA-FEE-Unicamp
e.mail: clesio@dca.fee.unicamp.br
Brazil

COMMISSION V

ABSTRACT

The paper addresses the problem of camera calibration and object location in robotics eye-in-hand applications. The proposed solution uses a modified functional model based on straight features for camera calibration and introduces Kalman filtering techniques for improvement of the results. A reduction of the computational effort of features extraction at image processing level is obtained by the feedback of estimated parameters of camera location. Results concerning precision and computational effort are presented and discussed.

KEY WORDS: Robot Vision, Machine Vision, Kalman Filtering, Camera Calibration, Space Resection, Eye-in-Hand, Object location.

1. INTRODUCTION

Current application of industrial robots, when used without external sensorial feedback, are limited by the uncertain models of the robots and unknown environments. Precision and flexibility of the whole system are increased when vision and others sensors, such as laser, sonar and tactile sensors are used.

Vision systems have been used mainly for recognizing, locating and inspecting stationary parts. However, visual information may be used to identify and locate objects or as a feedback to the robot control systems.

According to Weiss et al (1987), the actual robot geometry may be slightly different from the robot model and, therefore, the actual end-effector position may differ from the desired one. In order to solve this problem position and orientation (pose) of the manipulator end-effector obtained by vision sensors can be used as a feedback signal to control robot in real time (Feddema et al, 1991).

Vision systems can be introduced in the robot either by attaching a camera over the wrist (eye-in-hand) or in a remote position (eye-off-hand). In the eye-in-hand system the problem is the determination of camera location and orientation each time a robot movement is made. In the eye-off-hand configuration, otherwise, camera location and orientation is known and tracking and reconstructing the wrist position becomes the problem.

1.1 Camera Calibration

The problem of calculating camera position and orientation is called camera calibration or space resection. In Photogrammetry, besides the six position and orientation parameters, the problem of calibration involves additional inner parameters, which describe the internal camera geometry. In order to obtain the 6 external parameters, control points are used in most of photogrammetric and vision approaches. Given a set of image coordinates and corresponding world coordinates of control points, the well known collinearity equations and Least Squares Method can be used in order to get an optimal estimate for the parameters. This approach is iterative and requires linearization of the collinearity equations, which is time consuming and improper for real time applications. In order to avoid the computational cost caused by collinearity

linearization, some alternatives have been proposed which adopt linear models: Abdel-Aziz, Karara (1974), Lenz and Tsai (1988), Fischler and Bolles (1981).

Once the problem of parameters estimation is solved, remains the problem of feature extraction and correspondence of the control points in the image and object space, which is the bottleneck in the Machine Vision process. Most of the authors avoid this problem by considering correspondence as a foregoing step in their approaches.

Alternatives for features correspondence have been developed using, instead of points, more meaningful features, such as, straight lines, curved lines, rectangular shapes, junctions, etc.. Straight lines present advantages over other features considering that:

- . images of man-made environments are plenty in straight lines;
- . straight lines are easier to detect than point features and the correspondence problem becomes easier;
- . straight line parameters can be obtained with subpixel accuracy.

The use of alternative features has received more attention in recent years and more and more methods have been proposed: Masry (1980), Lugnani (1980), Tommaselli and Lugnani (1988), Mulawa and Mikhail (1988), Liu and Huang (1988a and 1988b), Salari and Jong (1990), Mitiche, Faugeras and Aggarwal (1989), Mitiche and Habelrih (1989), Dhome, Richetin, Lapresté and Rives (1989), Halarick (1989), Chen, Tseng and Lin (1989), Wang and Tsai (1990), Lee, Lu and Tsai (1990), Echigo (1990), Chen and Jiang (1991), Chen and Tsai (1991).

It is important to observe that methods for camera calibration in Machine Vision must take into account parameter estimation and error analysis in order to avoid unreliable solutions.

1.2 Filtering

Filtering techniques offer two great advantages when applied to the dynamic space resection problem in eye-in-hand systems: firstly, parameter estimation can be obtained using past observations without storing them; secondly, for each observation an state estimate is generated. This recursive approach can be used to feedback the feature extraction step, in order to reduce the search space both in image and Hough space and, therefore, to diminish computational effort.

A priori information about camera location and orientation (space resection) can be used to draw a polygon which probably contains the searched feature. A Hough transform can be limited to this subimage, reducing mathematical operations and generating a few candidate clusters. Using a filtering approach, parameter estimation is recursively improved for each new observation introduced and, the better is the camera location estimate, the smaller is the subimage to be processed.

Kalman Filtering is a suitable method to get optimal estimates for parameters in robotics eye-in-hand vision dynamic space resection.

Broida and Chellapa (1986) proposed an approach for the estimation of object motion parameters based on a sequence of noisy images. The movement of the body is modeled as a function of time and recursive filtering techniques are used to estimate the modeled parameters from the data measured in the images. In this approach image coordinates of two points in the body over a sequence of images are used as measurements.

A similar approach is adopted by Liang, Chang and Hackwood (1989) to address the problem of a vision-based robot manipulator system. In their method intrinsic (interior) and extrinsic (exterior) orientation parameters are recursively estimated based on several images of a single point.

1.3 Objectives

This paper presents a filtering and feature-based approach to space-resection (camera calibration). A mathematical model is developed which is based on straight features. This model is treated by Kalman filtering techniques, and it is shown that an iterative process between image processing and parameter estimation can be defined aiming a reduction on the computational effort.

2. GEOMETRIC MODEL FOR LINE CORRESPONDENCE

2.1 Object to image space transformation

There are several possible ways to define the sequence of transformations that relates two reference systems. In this paper, object to image space transformation is done by translating the origin of the object space until it reaches the image space origin and, then, rotating the resulting system around the resulting system axes. This transformation is described mathematically by equation (2.1.1) where X_c, Y_c, Z_c are the coordinates of the camera perspective center (image space origin), R is the rotation matrix, and Λ is a scale factor.

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \Lambda R \begin{bmatrix} X - X_c \\ Y - Y_c \\ Z - Z_c \end{bmatrix} \quad (2.1.1)$$

Let κ, ϕ and ω be the rotation sequence, the resulting rotation matrix R is given by:

$$\begin{bmatrix} \cos\phi \cos\kappa & \cos\omega \sin\kappa + \sin\omega \sin\phi \cos\kappa & \sin\omega \sin\kappa - \cos\omega \sin\phi \cos\kappa \\ -\cos\phi \sin\kappa & \cos\omega \cos\kappa - \sin\omega \sin\phi \sin\kappa & \sin\omega \cos\kappa + \cos\omega \sin\phi \sin\kappa \\ \sin\phi & -\sin\omega \cos\phi & \cos\omega \cos\phi \end{bmatrix} \quad (2.1.2)$$

2.2 The interpretation plane

The following concepts will be developed supposing that systematic errors in the image coordinates, such as symmetric radial distortion and horizontal scale factor, were previously eliminated and that the image coordinates were reduced to the principal point (image center). It will also be assumed that focal length is known and that the camera and object remain static during one image frame acquisition.

The *interpretation plane* is defined by a straight line in the image and the camera perspective center in image space reference system. Similarly, a straight line in the object space and the camera perspective center define a plane. This plane was called the *interpretation plane* by Barnard (1983).

Let $p_1 = (x_1, y_1, f)$ and $p_2 = (x_2, y_2, f)$ be two distinct image points defining a line ℓ ; the interpretation plane Φ associated with ℓ can be represented by its normal vector \vec{N}

$$\vec{N} = \vec{p}_1 \times \vec{p}_2 = \begin{bmatrix} f.(y_2 - y_1) \\ f.(x_1 - x_2) \\ x_2.y_1 - x_1.y_2 \end{bmatrix} \quad (2.2.1)$$

and the interpretation plane equation is given by

$$f.(y_2 - y_1).x + f.(x_1 - x_2).y + (x_2.y_1 - x_1.y_2).z = 0 \quad (2.2.2)$$

The representation based on the line described by two points is suitable for analytical photogrammetry, but not for digital photogrammetry and vision. The line representation in polar coordinates (θ - ρ) has advantages for applications in those areas mainly if Hough transform is used. For a description of Hough transform see Gonzalez and Wintz (1987).

Considering the parametric representation of a 2D straight line:

$$y = a.x + b \quad (2.2.3)$$

and its normal representation:

$$\cos\theta.x + \sin\theta.y - \rho = 0 \quad (2.2.4)$$

with:

$$a = -\cotan \theta = \tan \alpha \quad (2.2.5)$$

$$b = \rho/\sin\theta = -\rho/\cos\alpha \quad (2.2.6)$$

the equation of interpretation plane can be written as:

$$f.\cos\theta.x + f.\sin\theta.y - \rho.z = 0 \quad (2.2.7)$$

and the normal vector to the interpretation plane is:

$$\vec{N} = \begin{bmatrix} f.\cos\theta \\ f.\sin\theta \\ -\rho \end{bmatrix} \quad (2.2.8)$$

2.3 Measurement Model

The concept of interpretation plane can be useful in deriving models with particular features. Liu, Huang and Faugeras (1990) proposed a model which enables two steps for the camera parameters search. Tommaselli and Lugnani (1988) presented a model called "equivalent planes model" that is based on equivalence between parameters of the interpretation plane in object and image space. The approach to be developed in this paper is based on this model but eliminating the parameter (λ) which is a scale factor between parameters of the planes in object and image space.

Let \mathcal{L} be a straight line in the object space with the parametric representation:

$$\mathcal{L}: \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \end{bmatrix} + t. \begin{bmatrix} l \\ m \\ n \end{bmatrix} \quad (2.3.1)$$

where X_1, Y_1, Z_1 are the object coordinates of a known point in the line; l, m, n are the directions of the line and t is a parameter.

Let \vec{f} be the normal vector to the interpretation plane in the object space and defined by the vector product between the direction vector \vec{n} and the vector $(\vec{P}_C - \vec{C})$ (see Fig. 2.3.1).

Rotating the \vec{f} vector in order to compensate for the angular differences between object and image space reference systems, the normal vector \vec{N} and the rotated

vector \vec{f} become equivalent. This means that \vec{N} and \vec{f} are parallel but have different modulus (the value of components of the vectors \vec{N} and \vec{f} are multiple).

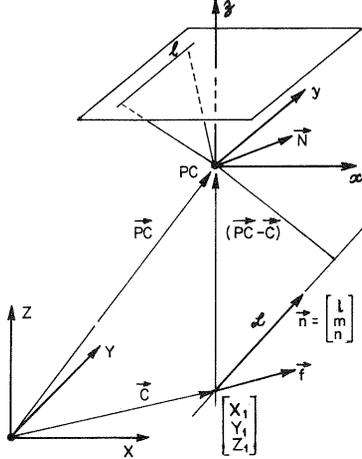


Figure 2.3.1 Normal vectors in object and image space.

The vector \vec{f} can be defined by the vector product:

$$\vec{f} = \vec{n} \times [\vec{PC} - \vec{C}] = \begin{bmatrix} 0 & n & -m \\ -n & 0 & 1 \\ m & -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} X_c - X_1 \\ Y_c - Y_1 \\ Z_c - Z_1 \end{bmatrix} \quad (2.3.2)$$

The relation between their components can be analytically described introducing a scale factor λ , which leads to:

$$\vec{N} = \lambda \cdot R \cdot \vec{f} \quad (2.3.3)$$

Combining equations 2.2.8, 2.3.2 and 2.3.3 the model can be stated more explicitly:

$$\begin{bmatrix} f \cdot \cos\theta \\ f \cdot \sin\theta \\ -\rho \end{bmatrix} = \lambda \cdot \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \cdot \begin{bmatrix} 0 & n & -m \\ -n & 0 & 1 \\ m & -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} X_c - X_1 \\ Y_c - Y_1 \\ Z_c - Z_1 \end{bmatrix} \quad (2.3.4)$$

In equation 2.3.4, the scale factor λ can be stated as the ratio of f and N modulus.

In order to eliminate λ in equations 2.3.4 the first and the third equations are divided by the second one, resulting:

$$\cot\theta = \frac{r_{11} \cdot f_x + r_{12} \cdot f_y + r_{13} \cdot f_z}{r_{21} \cdot f_x + r_{22} \cdot f_y + r_{23} \cdot f_z} \quad (2.3.5)$$

$$\frac{-\rho}{f \cdot \sin\theta} = \frac{r_{31} \cdot f_x + r_{32} \cdot f_y + r_{33} \cdot f_z}{r_{21} \cdot f_x + r_{22} \cdot f_y + r_{23} \cdot f_z}$$

Equations (2.3.5) can be rewritten as:

$$a = \frac{r_{11} \cdot f_x + r_{12} \cdot f_y + r_{13} \cdot f_z}{r_{21} \cdot f_x + r_{22} \cdot f_y + r_{23} \cdot f_z} \quad (2.3.6)$$

$$b = -f \cdot \frac{r_{31} \cdot f_x + r_{32} \cdot f_y + r_{33} \cdot f_z}{r_{21} \cdot f_x + r_{22} \cdot f_y + r_{23} \cdot f_z}$$

Equations (2.3.6) are indefinite when $\theta \cong 0^\circ$ (or $\alpha \cong 90^\circ$) and by this reason it can be used only for $45^\circ < \theta < 135^\circ$ or $225^\circ < \theta < 315^\circ$.

Defining a new representation for the straight line

$$x = a \cdot y + b \quad (2.3.7)$$

where: $a^* = -\tan\theta = \cot\alpha$ $b^* = \rho/\cos\theta = -\rho/\sin\alpha$

a complementary set of equations can be derived from 2.3.4 dividing the second and the third equations by the first one, which is applied for $315^\circ < \theta < 45^\circ$ or $135^\circ < \theta < 225^\circ$.

$$a^* = \frac{r_{21} \cdot f_x + r_{22} \cdot f_y + r_{23} \cdot f_z}{r_{11} \cdot f_x + r_{12} \cdot f_y + r_{13} \cdot f_z} \quad (2.3.8)$$

$$b^* = -f \cdot \frac{r_{31} \cdot f_x + r_{32} \cdot f_y + r_{33} \cdot f_z}{r_{11} \cdot f_x + r_{12} \cdot f_y + r_{13} \cdot f_z}$$

3. SYSTEM MODEL

3.1 Introduction

The problem of locating an object or calibrating a camera in Robot Vision is depicted in Figure 3.1.1. In this specific example it is assumed that the object is moving over a conveyor belt (a linear movement) and that the camera is attached to the robot wrist. In the eye-in-hand configuration the camera state vector is given by the kinematics of the robot plus a known transformation (a priori calibration) between wrist and camera coordinate systems. Due to cumulative errors in the robot joints the camera state vector can present some uncertainty.

3.2 Transform equations

The problem to be solved using eye-in-hand Vision can be better stated by defining the following homogeneous transformations:

- ${}^B T_S$ describes the station frame with respect to the base of the manipulator;
- ${}^S T_O$ describes the object frame with respect to station frame. The model of the object is specified in the object frame;
- ${}^O T_G$ describes the goal frame with respect to the object frame. The goal frame defines the position and orientation which must be reached by the manipulator wrist or an end effector;
- ${}^B T_W$ describes the wrist frame with respect to the base frame. This transformation is also known as the kinematics of the manipulator and is obtained by successive transformations over the links;
- ${}^W T_C$ describes the camera with respect to the wrist frame;
- ${}^C T_G$ describes the goal frame in camera coordinates;

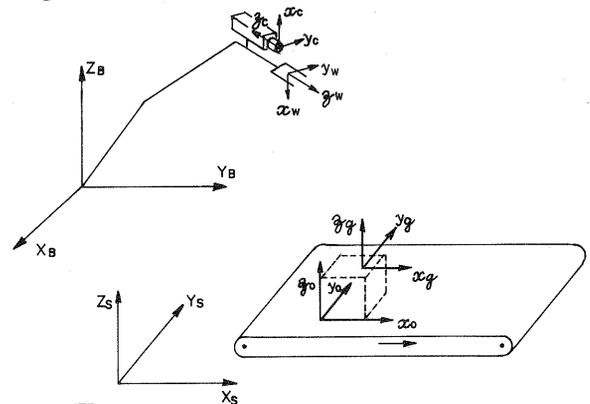


Figure 3.1.1 Frames and transforms in eye-in-hand problem.

The general problem of transformation between two reference frames can be stated by the transform graph (Paul, 1981) presented in Figure 3.1.2.

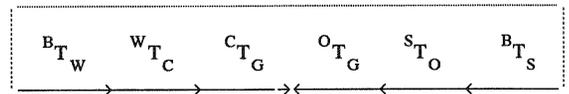


Figure 3.1.2 Transform graph.

In the eye-in-hand approach two cases can be stated:

• **first case:** the object is static and transforms ${}^W T_C$, ${}^O T_G$, ${}^S T_O$ and ${}^B T_S$ are known. Transform ${}^C T_G$ is computed using space resection techniques (in fact, transform ${}^C T_G^{-1}$). Then, transform ${}^B T_W$ can be obtained using the following transform equation:

$$B_{TW} = W_{TC} \cdot C_{TG} \cdot O_{TG}^{-1} \cdot S_{TO}^{-1} \cdot B_{TS}^{-1}$$

This procedure leads to an estimation of the kinematics of the manipulator using Vision and is in general known as **camera calibration**. This approach can be used to verify the real position and orientation of the wrist.

second case: it is supposed that transform B_{TW} is known (the kinematics model is used) and transform C_{TG} is computed. Once W_{TC} , O_{TG} , B_{TS} are also known, S_{TO} can be obtained using the following transform equation:

$$S_{TO} = O_{TG} \cdot C_{TG}^{-1} \cdot W_{TC}^{-1} \cdot B_{TW}^{-1} \cdot B_{TS}$$

The object must be located by the Vision system in order to be grasped by the wrist; this is called **object location** problem. In this case the accuracy of the position and orientation of the object is bounded by the error in the kinematics model.

4. FILTERING

4.1 Introduction

There are several parameters estimation methods, such as the well known Least Squares and the more recent Kalman Filtering. In Least Squares it is assumed that parameters are not time dependent and observation does not follow an a priori probability distribution.

In this paper, due to the dynamic nature of the problem, Kalman Filtering will be used for parameters estimation. In the next section a short review of Kalman Filtering is presented. A more detailed description can be obtained, for example, in Jazwinski (1970) and Miller & Leskiw (1987).

4.2 Kalman Filtering

Equation (4.2.1) describes a discrete and dynamic stochastic system:

$$x_{k+1} = \phi(x_k, t_{k+1}, t_k) + \Gamma(x_k, t_k) w_{k+1} \quad (4.2.1)$$

where: x_k is the n-vector state at t_k ;
 ϕ is a n-vectorial state transition function;
 Γ is a (nxr) matrix;
 w_k is a r-vector, white Gaussian sequence,
 $w_k \sim N(0, Q_k)$, usually called state transition noise;

Let z_k be the observation vector:

$$z_k = h(x_k, t_k) + n_k \quad k = 1, 2, \dots \quad (4.2.2)$$

where: z_k are the observations at t_k ;
 n_k is the vector of measurement noise,
 $n_k \sim N(0, R_k)$;

Equation (4.2.2) describes the *measurement model*.

When both the system model and the measurement model are linear, then:

$$\phi(x_k, t_{k+1}, t_k) = F(t_{k+1}, t_k) x_k \quad (4.2.3)$$

and: $h(x_k, t_k) = H(t_k) x_k \quad (4.2.4)$

In this case, to update the estimates:

$$\hat{x}_{k+1:k} = F_k \hat{x}_{k:k} \quad (4.2.5)$$

$$P_{k+1:k} = F_k P_{k:k} F_k^T + Q_k \quad (4.2.6)$$

4.2.1 Iterated Extended Kalman Filter (IEKF)

The original Kalman filter deals only with linear models. In order to use this approach to non-linear discrete models Taylor linearization is introduced. In the next steps the equations used to get a state estimate using IEKF are presented without further developments.

The IEKF approach is based on an iterator which is analyzed for each iteration in order to verify the convergence. This iterator is given by:

$$\hat{\eta}_{i+1} = \hat{x}_{k;k-1} + K_{k;\eta_i} (z_k - h(\eta_i, t_k) - M_{k;\eta_i} [\hat{x}_{k;k-1} - \hat{\eta}_i]) \quad (4.2.7)$$

where:

$K_{k;\eta_i}$ is the Kalman gain matrix at t_k using estimates for the state vector given by η_i . Kalman gain is expressed by:

$$K_{k;\eta_i} = P_{k;k-1} M_{k;\eta_i}^T (M_{k;\eta_i} P_{k;k-1} M_{k;\eta_i}^T + R_k^{-1})^{-1} \quad (4.2.8)$$

$M_{k;\eta_i}$ is the partial derivatives matrix of function h with respect to the elements of the state vector:

$$M_{k;\bar{x}_k} = \left[\frac{\partial h_i(\bar{x}_k, t_k)}{\partial x_m} \right] \quad (4.2.9)$$

$P_{k;k-1}$ is the predicted covariance matrix, which is obtained based on an update of the covariance matrix computed at t_{k-1} ;

$\hat{x}_{k;k-1}$ is the predicted state at t_k , based on measurements at t_{k-1} , computed using the linearized state transition function.

η_i is the iterator, which is an estimate for $\hat{x}_{k;k}$ at t_k . At the first iteration $\eta_1 = \hat{x}_{k;k-1}$, which is the predicted state estimate at t_k based on measurements taken at t_{k-1} . The result of the first iteration, η_2 , is used in the third iteration and this process is carried out until there is no further improvement in η_i . Thus, the final estimate for $\hat{x}_{k;k}$ is given by the last iterator η_i .

Since the state estimate has converged the filtered covariance matrix can be computed and it is given by:

$$P_{k:k} = (I - K_k M_k) P_{k;k-1} (I - K_k M_k)^T + K_k R_k K_k^T \quad (4.2.10)$$

When multiples uncorrelated observations are available at instant t_k it is possible to compute recursively the estimates for the state at instant t_k . This is accomplished introducing one observation at time and supposing that $F = I$ and $\Gamma = Q = 0$; that is, the state remains unchanged at t_k and its estimate and covariance matrix are recursively improved for each new observation introduced.

5. FILTERING BASED CALIBRATION AND OBJECT LOCATION

5.1 Introduction

The measurement model presented in Section 2.3 can be used in a filtering-based approach in order to reach a more accurate and flexible method of camera calibration and object location. For this purpose model equations must state explicitly the relations between straight lines parameters in image and object spaces and the state vector defined by camera (or object) orientation and position variables.

In the following sections the measurement equations for the filter are developed and is also discussed their relation to image feature extraction. It is considered that the object model is known, i.e., straight lines of the object have known parametric equations in the object reference frame and that straight lines parameters in image space are obtained using Hough transform.

5.2 Kalman Filtering measurement model

From Section 2.3 we have that the measurement model is defined by equations 2.3.6 for $45^\circ < \theta < 135^\circ$ or $225^\circ < \theta < 315^\circ$ and by equations 2.3.7 for $135^\circ < \theta < 45^\circ$ or $135^\circ < \theta < 225^\circ$.

Let x_k be the state vector at time t_k and G_1, G_2 be measurement model functions as defined by equations 2.3.6; hence,

$$G_1(x_k) = a = -\frac{r_{11}.f_x + r_{12}.f_y + r_{13}.f_z}{r_{21}.f_x + r_{22}.f_y + r_{23}.f_z} = \frac{v}{u}$$

$$G_2(x_k) = b = -f \frac{r_{31}.f_x + r_{32}.f_y + r_{33}.f_z}{r_{21}.f_x + r_{22}.f_y + r_{23}.f_z} = \frac{w}{u}$$

$$x_k = [\kappa, \phi, \omega, X_c, Y_c, Z_c]^T$$

The partial derivatives matrix of function G_1 with respect to the state vector is:

$$M_1 = \begin{bmatrix} \frac{\partial G_1}{\partial \kappa} & \frac{\partial G_1}{\partial \phi} & \frac{\partial G_1}{\partial \omega} & \frac{\partial G_1}{\partial X_c} & \frac{\partial G_1}{\partial Y_c} & \frac{\partial G_1}{\partial Z_c} \\ \frac{\partial G_2}{\partial \kappa} & \frac{\partial G_2}{\partial \phi} & \frac{\partial G_2}{\partial \omega} & \frac{\partial G_2}{\partial X_c} & \frac{\partial G_2}{\partial Y_c} & \frac{\partial G_2}{\partial Z_c} \end{bmatrix}$$

its elements are:

$$\frac{\partial G_1}{\partial \kappa} = -1 - \frac{v^2}{u^2} \quad \frac{\partial G_2}{\partial \kappa} = -f \cdot \frac{w \cdot v}{u^2}$$

$$\frac{\partial G_1}{\partial \phi} = \frac{-w \cdot u \cdot c\kappa - w \cdot v \cdot s\kappa}{u^2}$$

$$\frac{\partial G_2}{\partial \phi} = -f \cdot \frac{(c\phi \cdot f_x + s\omega \cdot s\phi \cdot f_y - c\omega \cdot s\phi \cdot f_z) \cdot u - s\kappa \cdot w^2}{u^2}$$

$$\frac{\partial G_1}{\partial \omega} = \frac{(-r_{13} \cdot f_y + r_{12} \cdot f_z) \cdot u - v \cdot (-r_{23} \cdot f_y + r_{22} \cdot f_z)}{u^2}$$

$$\frac{\partial G_2}{\partial \omega} = -f \cdot \frac{(-r_{33} \cdot f_y + r_{32} \cdot f_z) \cdot u - w \cdot (-r_{23} \cdot f_y + r_{22} \cdot f_z)}{u^2}$$

$$\frac{\partial G_1}{\partial X_c} = \frac{\frac{\partial v}{\partial X_c} u - v \frac{\partial u}{\partial X_c}}{u^2} \quad \frac{\partial G_2}{\partial X_c} = -f \cdot \frac{\frac{\partial w}{\partial X_c} u - w \frac{\partial u}{\partial X_c}}{u^2}$$

$$\frac{\partial G_1}{\partial Y_c} = \frac{\frac{\partial v}{\partial Y_c} u - v \frac{\partial u}{\partial Y_c}}{u^2} \quad \frac{\partial G_2}{\partial Y_c} = -f \cdot \frac{\frac{\partial w}{\partial Y_c} u - w \frac{\partial u}{\partial Y_c}}{u^2}$$

$$\frac{\partial G_1}{\partial Z_c} = \frac{\frac{\partial v}{\partial Z_c} u - v \frac{\partial u}{\partial Z_c}}{u^2} \quad \frac{\partial G_2}{\partial Z_c} = -f \cdot \frac{\frac{\partial w}{\partial Z_c} u - w \frac{\partial u}{\partial Z_c}}{u^2}$$

$$\frac{\partial u}{\partial X_c} = (-r_{22} \cdot n + r_{23} \cdot m) \quad \frac{\partial u}{\partial Y_c} = (r_{21} \cdot n - r_{23} \cdot l)$$

$$\frac{\partial v}{\partial X_c} = (-r_{12} \cdot n + r_{13} \cdot m) \quad \frac{\partial v}{\partial Y_c} = (r_{11} \cdot n - r_{13} \cdot l)$$

$$\frac{\partial w}{\partial X_c} = (-r_{32} \cdot n + r_{33} \cdot m) \quad \frac{\partial w}{\partial Y_c} = (r_{31} \cdot n - r_{33} \cdot l)$$

$$\frac{\partial u}{\partial Z_c} = (-r_{21} \cdot m + r_{22} \cdot l) \quad \frac{\partial w}{\partial Z_c} = (-r_{31} \cdot m + r_{32} \cdot l)$$

$$\frac{\partial v}{\partial Z_c} = (-r_{11} \cdot m + r_{12} \cdot l)$$

where: $c_i = \cos i$ and $s_i = \sin i$

The partial derivatives for the model defined by

equations 2.3.7 can be stated in a similar way.

5.2.2 Steps of the IEKF

At an instant t_k it is supposed that one image is acquired and that some of the straight lines detected in this image have known parametric equations in object space. In our particular measurement model, each straight line detected in the image space gives two equations. The state estimate can be improved with these observations at t_k . The greater the number of lines the better is the estimate. Even a single line will improve the state estimate, although this is a deficient configuration.

The IEKF is a recursive method for state estimation, which enables an observation to be processed once it becomes available. This feature makes feasible the definition of a recursive strategy in which the search for lines in the image is improved by better state estimates.

The steps to IEKF computation are described below:

1. In a initial instant the state vector $x_{1:0}$ and its covariance matrix $P_{1:0}$ are known. In the case of camera calibration these estimates are obtained using the robot kinematics. In the case of object location some probable position must be given;

2. Using predicted state estimate, at time t_k , the first line is searched in the edge subimage using Hough transform. This first straight line will provide measurements for a, b and an estimate for R (covariance matrix of the observations). A state estimate and a covariance matrix for the filtered state can be computed using these observations, with the following dimensions for vectors and matrix;

$${}^6 x_1 \quad {}^6 \delta x_1 \quad {}^2 M_6 \quad {}^2 R_2 \quad {}^6 P_6 \quad {}^6 K_2 \quad {}^2 z_1$$

3. The filtered state (and its covariance matrix) is used to reduce the search space for a new observation. A new subimage is defined (smaller than the first) and the second line in image space corresponding to a known 3D line in object space is found. The filtered state and covariance matrix obtained using j^{th} observation is considered as a predicted value for the $(j+1)^{\text{th}}$ observation and a new state estimate can be computed. This step can be repeated until all the available lines have been processed at t_k ;

4. Now, a prediction $\hat{x}_{k+1:k}$ and $P_{k+1:k}$ for the next instant t_{k+1} must be established using the available estimates and the system model. In the case of camera calibration the motion of the manipulator is known and the differential changes in the state are predicted using the Jacobian (Paul, 1981); an error estimate for this prediction is available, if joints uncertainties are known. In the case of object location usually the object is transported by a conveyor whose velocity is known, allowing state prediction.

The main advantage of this recursive approach is to enable the reduction of computational costs at image feature extraction level by the reduction of the search space.

5.3 Improvement of the image feature extraction process

In camera calibration the problem of correspondence, i.e., location of a known specific object feature in image space, has a high computational cost. Points and straight lines are the most common features in Vision Systems and for their detection template matching and Hough transform are used, respectively. For both procedures a search in the 2D image space is required and the greater is the image search space the higher will be the computational cost;

thus, so the reduction of the image search space is a way to increase the performance and efficiency of Vision Systems.

Estimates of camera state ($\hat{x}_{k+1:k}$ and $P_{k+1:k}$) can be used to predict the feature position in image space and the search limited to a window around this position. The prediction of the feature position in the image space can be done projecting object entities into image space using collinearity equations or even the measurement model defined in this paper. Prediction of any image feature position is bounded only by the quality of the state estimates. This way the window must be defined taking into account the covariance matrix of the predicted state estimate and the feature dimensions.

A recursive procedure in which sequential estimates are used to reduce the feature search space is depicted in Figure 5.3.1.

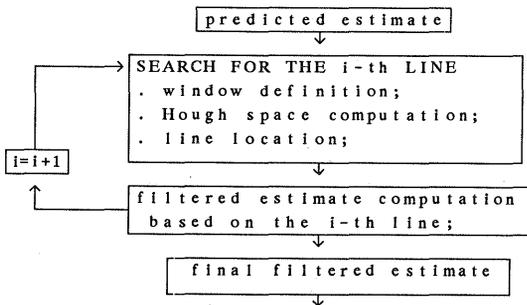


Figure 5.3.1 Recursive search procedure.

Other advantages of using predicted estimates for image features search are:

- the estimated feature length can be used to locate the most probable cluster in Hough space.
- fewer candidate clusters will be present in the Hough space.

6. RESULTS

6.1 Introduction

In the previous sections mathematical expressions relating straight features in object and image space and its treatment using Kalman filtering were presented. In this section results obtained from simulated data are presented and discussed.

Camera inner parameters (focal length, principal point, optical distortion) were supposed known and it was also assumed no movement of robot or object during image acquisition. The following parameters for the simulated camera were used: 15mm focal length, 10x10 mm² imaging area and 10x10 μm pixel size.

Once the exterior camera parameters (position and orientation) are established the image coordinates of object points (corners) were computed using collinearity equations. Random errors were introduced in these points which represent endpoints of a straight line; image line equations were finally computed from these pair of points.

Three sets of data became available:

- object straight lines parametric equations, assumed to be known from the object model;
- exterior camera parameters (camera vector state) and the associated error matrix;
- image lines equations and their covariance matrix. The covariance matrix of image lines were computed using covariance propagation.

6.2 Single Frame Calibration

It was supposed we had a single camera, static in space, observing a cube of 70mm. The base frame is coincident with the station frame and the object frame is 200mm far from the origin of the base frame. For the camera vector state of Table 6.2.1 the resulting

simulated image is presented in Figure 6.2.1.

Table 6.2.1 True Camera State and Predicted Values

	Camera State	Predicted State	State Error	Predicted Variance
κ	0.0	0.02	0.02	(0.02) ²
ϕ rad	0.0	-0.02	-0.02	(0.02) ²
ω	0.959931	0.939931	0.02	(0.02) ²
Xc	230	225	5	(5.) ²
Yc mm	-200	-204	4	(5.) ²
Zc	200	205	-5	(5.) ²

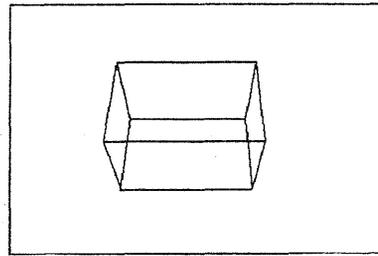


Figure 6.2.1 Simulated image of a cube.

The wireframe cube shown in Figure 6.2.1 can be described by twelve lines in image space, the correspondent object lines of which in base coordinates are known. Using the recursive approach stated in previous sections estimates for the camera vector are obtained. In Figure 6.2.2. graphics are presented showing the true errors and estimated standard deviations for rotation and translation variables. The true error is defined as the difference between the estimated and the true parameter value and the standard deviation is defined as the square root of the estimated variance.

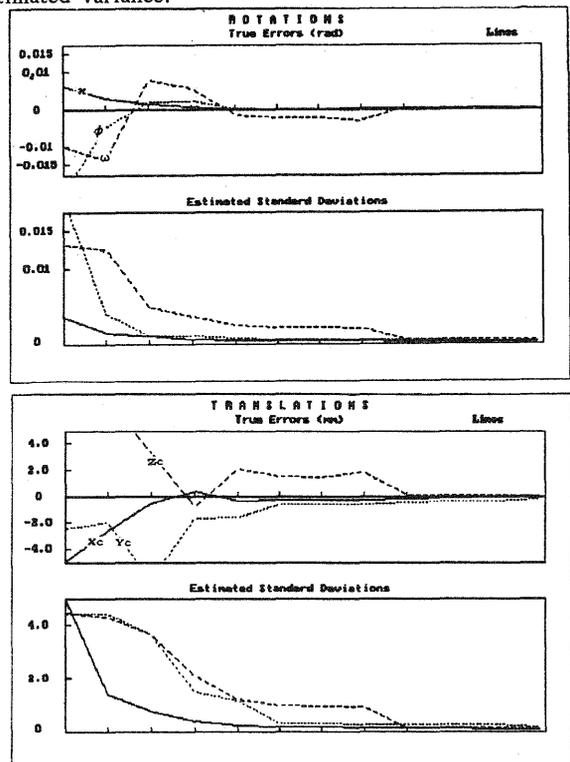


Figure 6.2.2 Error analysis in single frame calibration

From the analysis of the graphics in Figure 6.2.2 we can conclude:

- the filter has a strong convergence over the twelve lines. In fact, when the ninth feature was

introduced, the filter had already converged to the final values;
 only four lines are sufficient to obtain a good convergence;
 parameters κ and X_c converge before the others;
 a high accuracy estimation is reached with this approach; rotational and translational errors are smaller than 1' and 0.2 mm respectively.

6.3 Multi-frame Calibration

It was supposed we had a single camera, static in space, observing a 70 mm cube moving on a conveyor belt in X direction with a speed of 100 mm/s. A sequence of nine images is taken by the camera. In the first image the cube frame and base frame are coincident. Table 6.3.1 summarizes the camera vector state for this instant.

Table 6.3.1 True Camera State and Predicted Values

	Camera State	Predicted State	State Error	Predicted Variance
κ	0.0	0.01	0.01	$(0.01)^2$
ϕ rad	0.0	0.01	0.01	$(0.01)^2$
ω	0.959931	0.949931	-0.01	$(0.01)^2$
X_c	300	304	4.	$(4.)^2$
Y_c mm	-400	-396	4.	$(4.)^2$
Z_c	400	404	4.	$(4.)^2$

Knowing the conveyor belt speed and the sampling time, the object position can be computed and, then, the object-to-base frame transformations for each image can be stated. In order to simulate real environments, random perturbations were introduced in object position (1mm and 1° standard deviation in translation and rotation parameters, respectively).

For the same set of data, results obtained for a sequence of single frame calibration (a priori estimates for each image are not related to the former calibration) and a multi-frame calibration (a priori estimates are obtained from the filtered estimates of the former calibration) are presented in Figures 6.3.1 and 6.3.2, respectively.

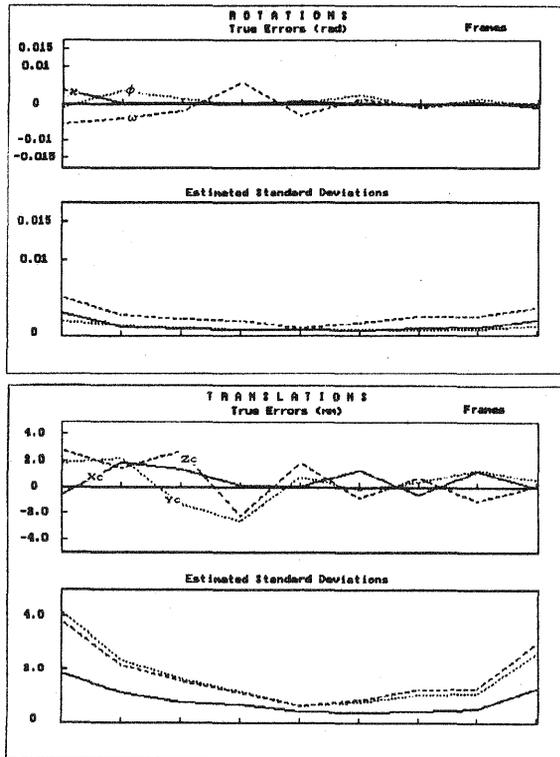


Figure 6.3.1 Results of a sequence of single frame calibrations

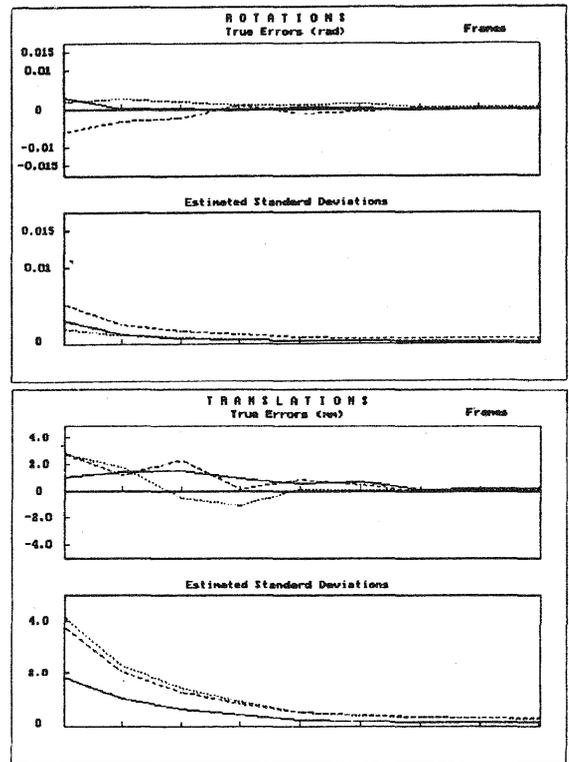


Figure 6.3.2 Results of a Multi-frame Calibration.

The improvements of the multi-frame calibration process can be seen comparing Figures 6.3.1 and 6.3.2. The final camera state estimation for the multi-frame calibration is within an accuracy of 2' for the rotations and 0.5 mm for the translations.

6.4 Reduction of the search window in feature extraction

In order to illustrate the reduction of the search space in the feature extraction level, the results related to the single frame calibration discussed in Section 6.2 are shown in Table 6.4.1.

Using the predicted estimates to define the first search space results in a rectangle of 56,072 pixels (1.16 x 3.6 mm), to be analysed. For the extraction of the 12th feature using the 11th estimate, the window is reduced to a rectangle of 0.031x 3.53 mm (a window of 3 pixels width), equivalent to 1,132 pixels.

Table 6.4.1 Reduction of the search window.

	Area mm ²	Num. of pixels	Window width mm	Feature Length
1	5.60	56,072	1.167	3.68
2	3.12	31,051	1.013	2.25
3	2.47	24,773	0.637	3.06
4	1.60	16,068	0.425	3.10
5	0.54	5,414	0.233	2.24
6	0.42	4,279	0.132	3.10
7	0.32	3,221	0.101	2.84
8	0.31	3,168	0.099	2.84
9	0.54	5,431	0.120	4.37
10	0.07	747	0.039	1.99
11	0.07	703	0.037	1.96
12	0.11	1,132	0.031	3.54

As can be seen from Table 6.4.1, the availability of better estimates for the camera state vector along the filter operation reduces the search window area. If only a priori estimates were used to search for the lines, 672,000 pixels (12 windows with 56,000 each one) should be analyzed, whereas using the recursive

procedure this value is reduced to 152,259 pixels. If instead of a window search, the whole image were used, 1,000,000 pixels should be analyzed. This simple simulation shows a reduction of 80% in the number of pixels to be analyzed; it is difficult to foresee the reduction in terms of floating point operations because only part of those pixels belong to edges and contribute to accumulation in Hough space.

7. CONCLUSIONS

We have presented a recursive approach for camera calibration and object location based on straight lines correspondences and state estimation using Kalman Filtering.

The derivation was presented of an explicit functional model which relates image and object straight lines. The Iterated Extended Kalman Filter was introduced and applied to the functional model aiming the estimation of camera to object transformation.

An iterative procedure was introduced for reduction of the search space in feature extraction level. It has been shown that this procedure enables a great optimization in time processing.

The proposed approach was tested using simulated data and feature extraction with sub-pixel accuracy. Results for single and multi-frame calibration were presented. The single frame calibration was used to show the filter convergence and the iterative reduction of the feature search windows, whereas the multi-frame calibration was used to show the filter convergence over several frames taken for different cube positions. It was shown that small noise in the predicted state vector does not affect the filter convergence.

Although a simple dynamic model of linear motion was used, it is expected that similar results arise for more complex models.

8. REFERENCES

- ABDEL-AZIZ, Y.I.; KARARA, H. M. Photogrammetric Potential of non-metric cameras. Univ. Illinois at Urbana-Champaign, Civil Eng. Studies, *Photogrammetry Series* 36, 1974.
- BARNARD, S. T. Interpreting Perspectives Images. *Artificial Intelligence*. N° 21, 1983.
- BROIDA, T.J.; CHELLAPPA, R. Estimation of Object Motion Parameters from Noisy Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol 8, N° 1, 1986.
- CHEN, S.Y.; TSAI, W.H. Determination of Robot Locations by Common Object Shapes. *IEEE Transactions on Robotics and Automation*. Vol. 7, N° 1, 1991.
- CHEN, Z.; TSENG, D.; LIN, J. A Simple Vision Algorithm for 3-d Position Determination using a Single Calibration Object. *Pattern Recognition*. Vol. 22, 1989.
- CHEN, W.; JIANG, B. C. 3-D Camera Calibration using Vanishing Point Concept. *Pattern Recognition*. Vol. 24, 1991.
- DHOME, M.; RICHTIN, M.; LAPRESTÉ, J.; RIVES, G. Determination of the Attitude of 3-D Objects from a Single perspective View. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol. 11, 1989.
- ECHIGO, T. A Camera Calibration Technique using Three Sets of Parallel Lines. *Machine Vision and Applications*. N° 3, 1990.
- FEDDEMA, J.T.; LEE, C.S.G.; MITCHELL, O.R. Weighted Selection of Image Features for Resolved Rate Visual Feedback Control. *IEEE Transactions on Robotics and Automation*. Vol. 7, N° 1, 1991.
- FISCHLER, A. M.; BOLLES, R. C. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM*. Vol. 24, N° 6, 1981.
- GONZALEZ, R. C.; WINTZ, P. *Digital Image Processing*. Addison-Wesley Publishing Company, 1987.
- HALARICK, R. M. Determining Camera Parameters from the Perspective Projection of a Rectangle. *Pattern Recognition*. Vol. 32, n° 3, 1989.
- JAZWINSKI, A.H. *Stochastic Processes and Filtering Theory*. New York. Academic Press, 1970.
- LEE, R.; LU, P.C; TSAI, W.H. Robot Location Using Single Views of Rectangular Shapes. *Photogrammetric Engineering and Remote Sensing*. Vol. 56, N° 2, 1990.
- LENZ, R. K.; TSAI, R. Y. Techniques for Calibration of the Scale Factor and Image Center for High Accuracy 3-D Machine Vision Metrology. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol. 10, n° 5, 1988.
- LIANG, P.; CHANG, Y. L.; HACKWOOD, S. Adaptive Self-Calibration of Vision-Based Robot Systems. *IEEE Transactions on Systems, Man, and Cybernetics*. Vol. 19, 1989.
- LIU, Y.; HUANG, T. S. Estimation of Rigid Body Motion Using Straight Line Correspondences. *Computer Vision, Graphics and Image Processing*. N° 43, 1988.
- LIU, Y.; HUANG, T. S. A Linear Algorithm for Motion Estimation Using Straight Line Correspondences. *Computer Vision, Graphics and Image Processing*. N° 44, 1988.
- LIU, Y.; HUANG, T. S.; FAUGERAS, O. D. Determination of Camera Location from 2-D to 3-D Line and Point Correspondences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol 12, N° 1, 1990.
- LUGNANI, J.B. Using Digital Entities as Control. PhD Thesis, Department of Surveying Engineering, UNB, 1980.
- MASRY, S. E. Digital Mapping using Entities: A New Concept. *Photogrammetric Engineering and Remote Sensing*. Vol. 48, 1981.
- MILLER, K.S.; LESKIW, D.M. *An Introduction to Kalman Filtering with Applications*. Robert E. Krieger Publishing Company, Malabar, Florida, 1987.
- MITICHE, A.; HABELRIH. Interpretation of Straight Line Correspondences using Angular Relations. *Pattern Recognition*. Vol. 22, 1989.
- MITICHE, A.; FAUGERAS, O.; AGGARWAL, J. K. Counting Straight Lines. *Computer Vision, Graphics and Image Processing*. N° 47, 1989.
- MULAWA, D. C.; MIKHAIL, E. M. Photogrammetric Treatment of Linear Features. *Proceedings of International Society for Photogrammetry and Remote Sensing*. Kyoto, 1988.
- PAUL, R. P. *Robot Manipulators: Mathematics, Programming and Control*. Cambridge, MA: MIT Press, 1981.
- SALARI, E.; JONG, C. A Method to calculate the Structure and Motion Parameters from Line Correspondences. *Pattern Recognition*. Vol. 23, N° 6, 1990.
- TOMMASELLI, A.M.G.; LUGNANI, J. B. An Alternative Mathematical Model to the Collinearity Equation using Straight Features. *Proceedings of International Society for Photogrammetry and Remote Sensing*. Kyoto, 1988.
- WANG, L.L.; TSAI, W.H. Computing Camera Parameters using Vanishing-Line Information from a Rectangular Parallelepiped. *Machine Vision and Applications*. N° 3, 1990.
- WEISS, L. E.; SANDERSON, A. C.; NEUMAN, C. P. Dynamic Sensor-Based Control of Robots with Visual Feedback. *IEEE Transactions on Robotics and Automation*. Vol RA-3, N° 5, 1987.