A ROBUST METHOD FOR CAMERA CALIBRATION AND 3-D RECONSTRUCTION FOR STEREO VISION SYSTEMS

Lenildo C. Silva, Mariane R. Petraglia, Antonio Petraglia

PEE/COPPE-EE, Federal University of Rio de Janeiro
CP 68504, CEP 21945-970, Rio de Janeiro, RJ, Brazil
{lenildo, mariane, petra}@pads.ufrj.br

ABSTRACT

This work presents a camera calibration procedure for a stereo vision system to be applied in visual inspection activities involving the three-dimensional reconstruction of a scene. The presented procedure encompasses a robust method developed for solving the non-linear least-squares problems encountered, in order to obtain the global solution, and hence achieve the smallest error in estimating the parameters of the exterior orientation of the camera system. Stemming from the theoretical analysis of the camera orientation problem, the development of the robust method entailed a combination of optimization techniques. Experiments with real images were performed to verify the robustness of the proposed approach.

1. INTRODUCTION

One of the most interesting applications of stereo vision is 3-D reconstruction, when three-dimensional information of a point in a scene is recovered from two or more views of that scene. This information can be presented in different forms (e.g. 3-D coordinates, discrete measures, angles, etc), that can be estimated by applying stereophotogrammetric techniques. Stereophotogrammetry includes a set of techniques from which inferences can be made about the position and orientation of a 3-D object, given its 2-D projections in each image of a stereo pair.

Camera calibration is an important stage in the 3-D reconstruction procedure. It comprises two main tasks: the estimation of the parameters that determine the relation between the scene and its projection in the camera plane, and the estimation of the internal parameters of the camera, referred respectively as exterior and interior orientation parameters. The exterior orientation procedure consists of estimating the rotation and translation that relates the camera reference frame to the world (or object) reference frame, and can be represented by a set of six parameters: three translation parameters and three rotation angles. The interior orientation procedure consists of estimating a set of internal parameters that control the projection of a 3-D point in the 2-D plane, independently on the position and orientation of the observed scene. The parameters to be estimated are the coordinates of the principal point, the parameters related to the focal distance, and the parameters that model the geometric distortion introduced by the camera lenses.

Several methodologies have been applied to determine the exterior and interior orientation parameters, using linear and/or non-linear methods. Linear methods have the advantage of low computational cost, but are very sensitive to noise [1]. On the other hand, non-linear methods require an approximate initial estimate to guarantee global convergence. In [2] the exterior and interior orientation parameters are estimated linearly, applying projective geometry concepts. In [3] a linear approach for the solution of the exterior orientation problem using a series of linear combinations and constraints is developed to estimate the unknown rotation and translation parameters. In [4] a set of analytical photogrammetric formulas is presented for the solution of the orientation problems described by non-linear least-squares methods, and solved iteratively from a given approximate initial solution. In [1] the exterior and interior orientation parameters are estimated by a non-linear procedure that also considers distortion effects introduced by the camera lenses. In [5] a least-squares methodology is proposed for the on-line determination of the exterior orientation using a backprojection algorithm and weak-perspective scheme for the initial estimation.

In the present paper a robust method is advanced for the estimation of the exterior orientation parameters, incorporating various minimization techniques based in non-linear least-squares algorithms. The initial estimate is provided by a linear method derived from the projective geometry theory.

2. 3-D RECONSTRUCTION

The relation between a given point $[x\, y\, z]^{\mathrm{T}}$ in the world reference frame and the corresponding point $[p\, q\, s]^{\mathrm{T}}$ in the camera reference frame can be established by a translation followed by a sequence of three rotations:

$$\begin{bmatrix} p \\ q \\ s \end{bmatrix} = \mathbf{R}(\omega,\phi,\kappa) \begin{bmatrix} x-x_0 \\ y-y_0 \\ z-z_0 \end{bmatrix} \tag{1}$$

where $\mathbf{t} = [x_0\, y_0\, z_0]^{\mathrm{T}}$ is the translation vector, and $\mathbf{R}(\omega,\phi,\kappa)$ is the rotation matrix. From this representation of the 3-D point in the camera reference frame, it is obtained the respective 2-D projection in the image plane, whose coordinates can be written as

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} + \frac{f}{s} \begin{bmatrix} p \\ q \end{bmatrix} \tag{2}$$

where f is the distance of the image plane to the camera lens, being related to the focal distance, and $[u_0\ v_0]^{\mathrm{T}}$ are the coordinates of the principal point. From (1) and (2) we obtain

$$\frac{x-x_0}{z-z_0} = \frac{r_{11}(u-u_0)+r_{21}(v-v_0)+r_{31}f}{r_{13}(u-u_0)+r_{23}(v-v_0)+r_{33}f} \tag{3}$$

$$\frac{y-y_0}{z-z_0} = \frac{r_{12}(u-u_0)+r_{22}(v-v_0)+r_{32}f}{r_{13}(u-u_0)+r_{23}(v-v_0)+r_{33}f} \tag{4}$$

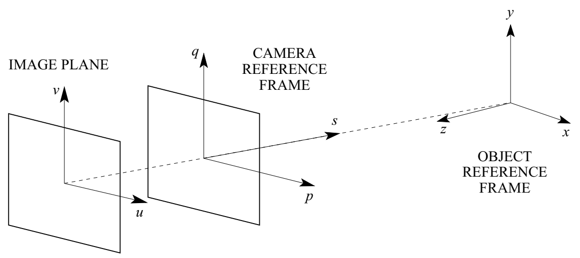


Figure 1: World and camera reference frames.

These equations show that the relation between the 3-D and 2-D coordinates is a function of f, u_0, v_0, x_0, y_0, z_0, ω, ϕ, κ, which can be determined through a non-linear least-squares technique.

3. NON-LINEAR LEAST-SQUARES METHOD

To estimate the exterior orientation, N 3-D points $[x_n\ y_n\ z_n]^{\mathrm{T}}$, $n = 1, \ldots, N$, with known positions in the world reference frame, are employed with the purpose of finding the unknown rotation and translation parameters that position the camera reference frame on the world reference frame [4, 6]. This can be formulated as a non-linear least-squares problem, which can be solved by starting from an approximate initial solution around which a linear model is produced, and then adjusting iteratively partial solutions until a given stopping criterion is achieved.

Adopting the cost function $F = \varepsilon^{\mathrm{T}}\varepsilon$, where $\varepsilon = \gamma^* - \gamma$ is the error vector, $\gamma^* = [u_1\ v_1\ \cdots\ u_N\ v_N]^{\mathrm{T}}$ contains the known image points and γ is the estimate of γ^*, being a function of the exterior orientation parameters $\beta = [x_0\ y_0\ z_0\ \omega\ \phi\ \kappa]^{\mathrm{T}}$, then these parameters can be obtained through the recursion

$$\beta^{\ell+1} = \beta^{\ell} + \Delta\beta^{\ell} \tag{5}$$

whose linearization around the solution at iteration ℓ yields

$$\mathbf{A}^{\ell}\Delta\beta^{\ell} = \varepsilon^{\ell} \tag{6}$$

where $\mathbf{A}^{\ell}$ is the $N \times N$ Jacobian matrix [4], and ε^{ℓ} is the error vector with image points estimated through Eqs. (1) and (2) using the parameters β^{ℓ}. The least-square solution of Eq. 6 is

$$\Delta\beta^{\ell} = [(\mathbf{A}^{\ell})^{\mathrm{T}}\mathbf{A}^{\ell}]^{-1}(\mathbf{A}^{\ell})^{\mathrm{T}}(\gamma^* - \gamma^{\ell}) \tag{7}$$

The interior orientation is determined by the camera constant f (a known parameter related to the focal distance), by the coordinates of the principal point $[u_0\ v_0]^{\mathrm{T}}$ and by the lens distortion characteristics.

If a good initial estimate can be provided, then the convergence of the above least-squares algorithm will be rapidly reached, although convergence cannot be guaranteed. In Eq. (7) the estimation of $\Delta\beta$ requires the inversion of the matrix $\mathbf{A}^{\mathrm{T}}\mathbf{A}$, which is usually ill-conditioned. To eliminate this problem, the Levenberg–Marquardt method [7, 8] was here applied, by adding a factor λ to the diagonal elements of $\mathbf{A}^{\mathrm{T}}\mathbf{A}$, so that a feasible condition number for this matrix could be achieved. The value of λ was adjusted at every iteration to accommodate variations of the cost function $F = \varepsilon^{\mathrm{T}}\varepsilon$, as follows. At each iteration, $F^{\ell+1}$ was evaluated and compared with F^{ℓ}. If $F^{\ell+1} \geq F^{\ell}$ the value of λ_{ℓ} was increased, and the iteration was again executed until $F^{\ell+1} < F^{\ell}$ was reached. Accordingly, Eq. (7) was modified as:

$$\Delta\beta^{\ell} = [(\mathbf{A}^{\mathrm{T}})^{\ell}\mathbf{A}^{\ell} + \lambda_{\ell}\mathbf{I}]^{-1}(\mathbf{A}^{\mathrm{T}})^{\ell}(\gamma^* - \gamma^{\ell}) \tag{8}$$

The vector $\Delta\beta$ is composed of variations of the translation parameters x_0, y_0 and z_0, given in meters, and of the rotation parameters ω, ϕ, and κ, given in radians. Such differences in scale and magnitude can again cause ill-conditioning in the least-squares procedure. This difficulty was alleviated by using the Sparse Levenberg–Marquardt approach [7], to partition the parameter vector into two sub-vectors β_t and β_r, containing the translation and rotation parameters, respectively. The updating vector $\Delta\beta$ was also partitioned into two sub-vectors, $\Delta\beta_t$ and $\Delta\beta_r$, computed separately at each iteration.

Another important change was introduced in Eq. (5). Since large values of the elements of $\Delta\beta$ might lead to local minima or even to divergence, a factor α_{ℓ} was introduced, such that

$$\beta^{\ell+1} = \beta^{\ell} + \alpha_{\ell}\Delta\beta^{\ell} \tag{9}$$

This factor could be held constant in all iterations, but better results were obtained by using a variable step size, modified at each iteration according to the Armijo rules [8], that is:

1. $F^{\ell+1} \leq F^{\ell} + \rho_1\alpha_{\ell}(\mathbf{g}^{\mathrm{T}})^{\ell}\Delta\beta$, for some $0 < \rho_1 \leq 1$
2. $(\mathbf{g}^{\mathrm{T}})^{\ell+1}\Delta\beta^{\ell} \leq \rho_2(\mathbf{g}^{\mathrm{T}})^{\ell}\Delta\beta$, for some $\rho_1 < \rho_2 \leq 1$

where $\mathbf{g} = 2\mathbf{A}^{\mathrm{T}}\varepsilon$ is the gradient vector of F. The value of α_{ℓ} was chosen according to $\rho_1\alpha_{\ell} \leq \alpha_{\ell+1} \leq \rho_2\alpha_{\ell}$.

4. LINEAR METHOD

To guarantee a fast convergence of the exterior orientation computation, an approximate initial solution of the non-linear least-squares algorithm is needed. This can be accomplished by using a linear method derived from the projective geometry theory. Although usually sensitive to measurement noise, as shown later in Section 6.1, this linear method provides good initial estimates for the exterior and interior orientation parameters, also called extrinsic and intrinsic parameters, respectively.

A point in the 3-D space can be linearly related to its 2-D projection in the image plane as [2]

$$\mathbf{m} = \mathbf{PM} \tag{10}$$

where $\mathbf{m} = [U, V, S]^{\mathrm{T}}$ and $\mathbf{M} = [X, Y, Z, T]^{\mathrm{T}}$ are, respectively, the projective coordinates of the 2-D and 3-D points. Eq. (10) is termed projective, because it is defined up to a scale factor (S and T in the above vectors). The matrix $\mathbf{P}$ is the so-called perspective projection matrix, and contains implicitly all the extrinsic and intrinsic parameters. The general form of $\mathbf{P}$ is

$$\mathbf{P} = \begin{bmatrix} \alpha_u\mathbf{r}_1 + u_0\mathbf{r}_3 & \alpha_u t_x + u_0 t_z \\ \alpha_v\mathbf{r}_2 + v_0\mathbf{r}_3 & \alpha_v t_y + v_0 t_z \\ \mathbf{r}_3 & t_z \end{bmatrix} \tag{11}$$

The six extrinsic parameters are then obtained from the translation vector $\mathbf{t} = [t_x\ t_y\ t_z]^{\mathrm{T}}$ and the rotation matrix $\mathbf{R} = [\mathbf{r}_1\ \mathbf{r}_2\ \mathbf{r}_3]^{\mathrm{T}}$, and used as a starting point of the least-squares recursions for the computation of the exterior orientation. The intrinsic parameters α_u, α_v (the focal distance in horizontal and vertical pixels), u_0 and v_0 are also needed for the exterior orientation computation.

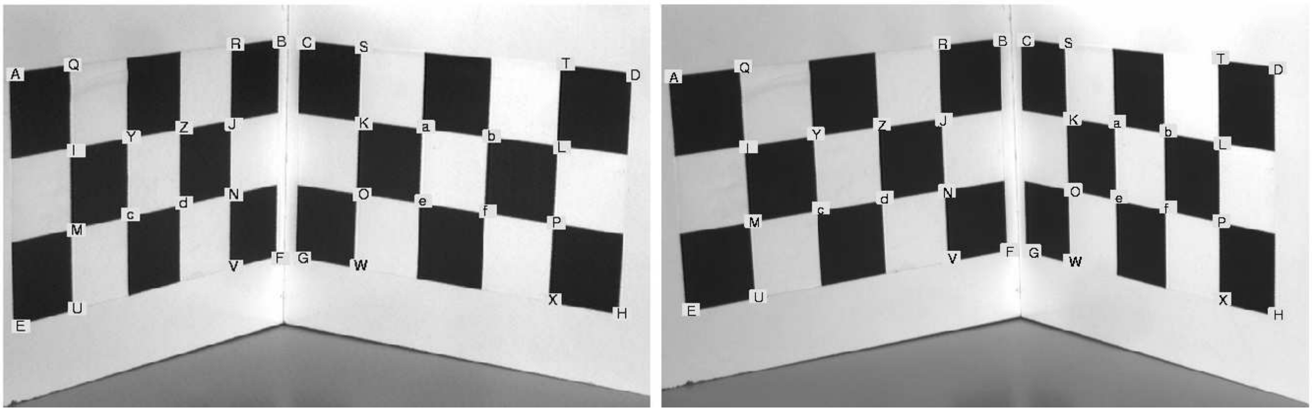


Figure 2: Stereo image pair used in the experiments.

Table 1: Experimental results for different numbers of points N.

SEGMENT	REAL	N = 32		N = 24		N = 16		N = 12		N = 8	
		ESTIM.	ERROR	ESTIM.	ERROR	ESTIM.	ERROR	ESTIM.	ERROR	ESTIM.	ERROR
$\overline{AB}$	0.250000	0.248528	0.589 %	0.248227	0.709 %	0.248371	0.652 %	0.248158	0.737 %	0.249280	0.288 %
$\overline{CG}$	0.150000	0.150711	0.474 %	0.150676	0.451 %	0.150604	0.403 %	0.150348	0.232 %	0.150485	0.323 %
$\overline{AD}$	0.367696	0.365426	0.617 %	0.365520	0.592 %	0.366815	0.239 %	0.366826	0.236 %	0.367693	0.001 %
$\overline{AC}$	0.260192	0.259616	0.222 %	0.259334	0.330 %	0.259318	0.336 %	0.259296	0.344 %	0.260563	0.143 %
$\overline{IJ}$	0.150000	0.150001	0.001 %	0.149972	0.018 %	0.150171	0.114 %	0.150056	0.038 %	0.150541	0.361 %
$\overline{IM}$	0.050000	0.049368	1.265 %	0.049438	1.124 %	0.049689	0.623 %	0.049559	0.883 %	0.049268	1.465 %
$\overline{IL}$	0.296985	0.297575	0.199 %	0.297779	0.268 %	0.297993	0.339 %	0.298160	0.396 %	0.299027	0.688 %
$\overline{IK}$	0.218403	0.221105	1.237 %	0.221138	1.252 %	0.220537	0.977 %	0.221117	1.243 %	0.222364	1.813 %
$\overline{SW}$	0.150000	0.150334	0.223 %	0.150388	0.259 %	0.150683	0.456 %	0.150452	0.301 %	0.150414	0.276 %
$\overline{TX}$	0.150000	0.151667	1.111 %	0.152098	1.399 %	0.153756	2.504 %	0.153531	2.354 %	0.152791	1.861 %
$\overline{QT}$	0.296985	0.298202	0.410 %	0.298064	0.363 %	0.298585	0.539 %	0.298513	0.515 %	0.299620	0.887 %
$\overline{QS}$	0.218403	0.217975	0.196 %	0.217751	0.299 %	0.217140	0.578 %	0.217568	0.383 %	0.218968	0.258 %
$\overline{ab}$	0.050000	0.049500	1.001 %	0.049532	0.935 %	0.049654	0.691 %	0.049705	0.591 %	0.049929	0.142 %
$\overline{Yc}$	0.050000	0.050357	0.713 %	0.050407	0.813 %	0.050603	1.206 %	0.050451	0.902 %	0.050222	0.444 %
MEAN		0.589 %		0.629 %		0.689 %		0.653 %		0.639 %	
STANDARD DEVIATION		0.420 %		0.421 %		0.595 %		0.587 %		0.628 %	

5. STEREO TRIANGULATION

In what follows, the subscripts L and R refer, respectively, to left and right parameters of a stereo pair. A stereo triangulation scheme is applied to estimate the 3-D coordinates of a point $[x\ y\ z]^{\mathrm{T}}$, from the 2-D coordinates of its projection on each image of a stereo pair ($[u_L\ v_L]^{\mathrm{T}}$ and $[u_R\ v_R]^{\mathrm{T}}$). The exterior and interior parameters computed for each image are the translation vectors $[x_L\ y_L\ z_L]^{\mathrm{T}}$ and $[x_R\ y_R\ z_R]^{\mathrm{T}}$, the rotation matrices $\mathbf{R}_L$ and $\mathbf{R}_R$, and the parameters f_L and f_R related to the focal distance. The 3-D coordinates $[x\ y\ z]^{\mathrm{T}}$ are estimated through one of the following equations:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_L \\ y_L \\ z_L \end{bmatrix} + \lambda_L \mathbf{R}_L \begin{bmatrix} u_L \\ v_L \\ f_L \end{bmatrix} \tag{12}$$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_R \\ y_R \\ z_R \end{bmatrix} + \lambda_R \mathbf{R}_R \begin{bmatrix} u_R \\ v_R \\ f_R \end{bmatrix} \tag{13}$$

It involves the estimation of the minimizing parameters λ_L and λ_R, found by performing the minimization of the squared difference

$$\varepsilon^2 = \left\| \begin{bmatrix} x_L \\ y_L \\ z_L \end{bmatrix} + \lambda_L \mathbf{R}_L \begin{bmatrix} u_L \\ v_L \\ f_L \end{bmatrix} - \begin{bmatrix} x_R \\ y_R \\ z_R \end{bmatrix} - \lambda_R \mathbf{R}_R \begin{bmatrix} u_R \\ v_R \\ f_R \end{bmatrix} \right\|^2 \tag{14}$$

6. EXPERIMENTAL VERIFICATION

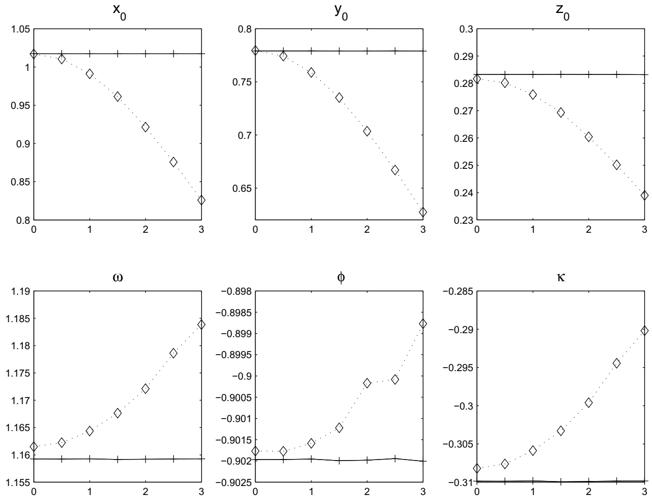
This section presents an illustrative application of the proposed camera calibration and 3-D reconstruction methods. The stereo pair of images used in the experiments is depicted in Fig. 2. The chosen 32 points are labeled by letters. The experiments entailed the estimation of some distances in Fig. 2, from the estimated 3-D coordinates of the corresponding delimiting vertices.

The measurement results are listed in Table 1, where the real and estimated dimensions are compared, and the relative error for each measure is computed. The results are shown in terms of the number of points N selected from the set $\{8, 12, 16, 24, 32\}$ during the camera calibration setup. The mean and the standard deviation of the relative error for each set of experiments are given at the bottom of Table 1. It can be observed that the mean error is smaller than 0.69%, and does not change significantly as the number of points N is increased. Hence, a small value of N can be chosen in the calibration approach proposed in this paper, consequently saving computational costs, without compromising the 3-D reconstruction performance.

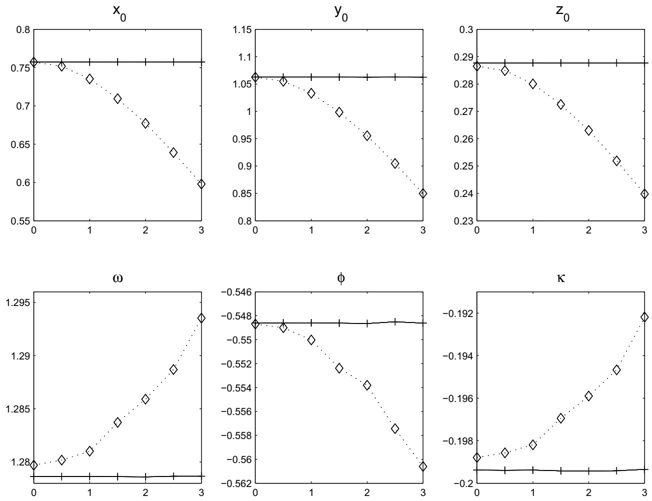
6.1 Effects of Noise in the Extrinsic Parameters

To verify the robustness of the proposed least-squares implementation, a Gaussian noise having zero mean and standard deviation σ_n varying from 0 (no noise) to 3 was added to the 2-D coordinates of the image points, to model measurement noise, and the resulting effects in the estimated exterior orientation parameters were observed.

About 1,000 independent simulations were carried out for each value of σ_n, and the mean value of each set of simulations was computed. These simulations were compared with the ones obtained using the linear method described in Section 4. The results obtained for each parameter are plotted in Figs. 3(a) and (b), for the left and right cameras, respectively. The translation parameters (x_0, y_0 and z_0) are expressed in meters, and the rotation parameters (ω, ϕ and κ) are expressed in radians. In all plots the horizontal axes indicate the values of the noise standard deviation. The curves with crosses refer to the proposed non-linear least-squares method, while the curves with diamonds refer to the linear method. From these plots, we conclude that the proposed algorithm is substantially less influenced by measurement noise than is the linear method, since the estimates obtained with the former remain very close to the actual values, regardless the increase in the noise variance. With the linear algorithm, on the other hand, as the noise variance increases, so do the estimation errors.



(a) left image



(b) right image

Figure 3: Effects of Gaussian noise in the extrinsic parameters estimation (see text).

7. CONCLUSIONS

A robust approach for camera calibration and 3-D reconstruction was presented. Used in the exterior orientation stage of stereo camera systems, the method lead to accurate and fast parameter estimation. The minimization techniques incorporated to the camera calibration process enabled global convergence of the derived non-linear least-squares algorithm. The small errors achieved in the 3-D reconstruction experiments verified the robustness of the proposed approach. The errors remained small regardless of the number of points used in the calibration procedure, showing the computational efficiency of the algorithm. The effectiveness of the robust method was also corroborated experimentally by analyzing the noise effects in the estimated parameters.

REFERENCES

[1] J. Weng, P. Cohen, and M. Herniou, "Camera Calibration with Distortion Models and Accuracy Evaluation," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 14, no. 10, pp. 965–980, Oct. 1992.

[2] O. Faugeras, *Three-Dimensional Computer Vision: a Geometric Viewpoint*, The MIT Press, Cambridge, USA, 1993.

[3] P. D. Fiore, "Efficient Linear Solution of Exterior Orientation," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 23, no. 2, pp. 140–148, Feb. 2001.

[4] R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision*, vol. 2, Addison-Wesley Publishing Company, 1993.

[5] C.-P. Lu, E. Mjolsness, and G. D. Hager, "Online Computation of Exterior Orientation with Application to Hand-Eye Calibration," *Mathl. Comput. Modeling*, vol. 24, no. 5/6, pp. 121–143, 1996.

[6] L. C. Silva, A. Petraglia, and M. R. Petraglia, "Stereo Vision System for Real Time Inspection and 3D Reconstruction," in *IEEE International Symposium on Signal Processing and Its Applications – ISSPA2004*, Paris, France, July 1–4, 2003.

[7] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2000.

[8] A. Jennings and J.J. McKeown, *Matrix Computation*, John-Wiley & Sons, Inc., 2 edition, 1993.