Appendices

Appendix A

Camera Calibration: Mathematical Developments

A.1 Introduction

The starting point of the development is the examination of the constraints on the seven intrinsic parameters of the camera (detailed in Section 2.4), provided by observing a planar object, called the *model plane*.

A.2 Homography between the Model Plane and its Image

Let (u, v, and w) be an orthonormal coordinate system attached to the observed model plane of Fig. A.1. We can rewrite equation (2.5), if we don't take into account distortion d at this point, as:

$$\begin{pmatrix} x_{pix} \\ y_{pix} \\ 1 \end{pmatrix} = \lambda A \begin{pmatrix} R & t \end{pmatrix} \begin{pmatrix} u \\ v \\ w \\ 1 \end{pmatrix} = \lambda A \begin{pmatrix} r_1 & r_2 & r_3 & t \end{pmatrix} \begin{pmatrix} u \\ v \\ w \\ 1 \end{pmatrix} \tag{A.1}$$

where R and t, called the extrinsic parameters, are the rotation and translation relating the model plane coordinate system to the WCS (r_1, r_2 and r_3 are chosen orthonormal, as such rotations always exist)

Without loss of generality, we assume that the model plane is on $w = 0$ and contains axes u and v (cf. Fig. A.1). From (A.1) we have:

$$\begin{pmatrix} x_{pix} \\ y_{pix} \\ 1 \end{pmatrix} = \lambda A \begin{pmatrix} r_1 & r_2 & t \end{pmatrix} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = H \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \tag{A.2}$$

where H is a 3x3 matrix, called *homography*.

We will use the abbreviation A^{-T} for $\left(A^T\right)^{-1}$ or $\left(A^{-1}\right)^T$.

A.3 Constraints on the Intrinsic Parameters

Given an image of the model plane, it was shown in [11] that a homography between the model plane and its image can be estimated (see [11] for further mathematical details). Let us denote it by $H = \begin{pmatrix} h_1 & h_2 & h_3 \end{pmatrix}$. From (A.2) we have:

$$\begin{pmatrix} h_1 & h_2 & h_3 \end{pmatrix} = \lambda A \begin{pmatrix} r_1 & r_2 & t \end{pmatrix} \tag{A.3}$$

Using the property that r_1 and r_2 are orthonormal, we have:

$$\begin{array}{rcl} h_1^T A^{-T} A^{-1} h_2 & = & 0 \\ h_1^T A^{-T} A^{-1} h_1 & = & h_2^T A^{-T} A^{-1} h_2 \end{array} \tag{A.4}$$

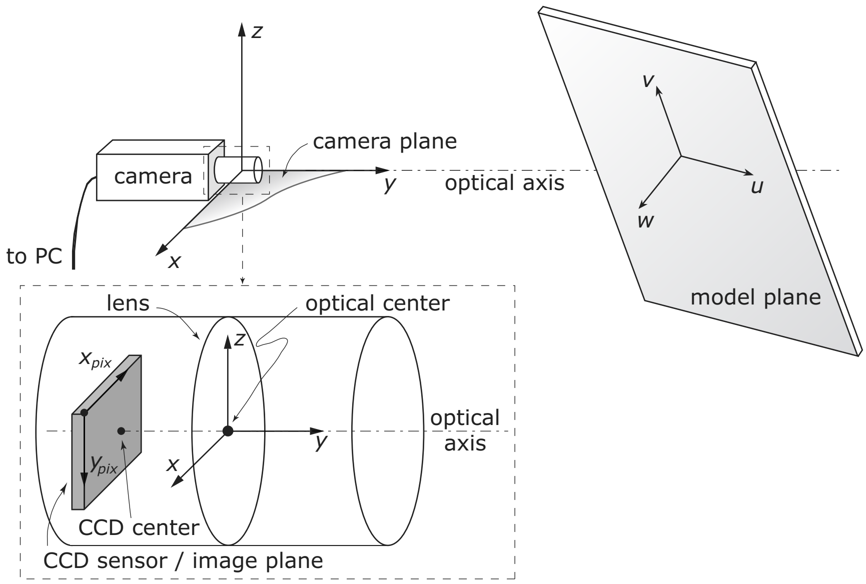

Fig. A.1. Camera calibration. Configuration of the coordinate systems.

These are the two constraints on the intrinsic parameters of the camera, given one homography H.

A.4 Closed-form Solution

One can find an analytical solution to the system described by (A.4). Let

$$B = A^{-T}A^{-1} = \begin{pmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{pmatrix} \tag{A.5}$$

By definition, B is symmetric, so that we have:

$$h_i^T B h_j = g_{ij} b \tag{A.6}$$

with

$$\begin{aligned} g_{ij} &= (h_{i1}h_{j1} \quad h_{i1}h_{j2} + h_{i2}h_{j1} \quad h_{i2}h_{j2} \quad h_{i3}h_{j1} + h_{i1}h_{j3} \\ &\qquad h_{i3}h_{j2} + h_{i2}h_{j3} \quad h_{i3}h_{j3}) \\ b &= (B_{11} \quad B_{12} \quad B_{22} \quad B_{13} \quad B_{23} \quad B_{33}) \end{aligned} \tag{A.7}$$

Therefore, the two fundamental constraints (A.4), from a given homography, can be rewritten as two homogeneous equations in b:

$$\begin{pmatrix} g_{12} \\ g_{11} - g_{22} \end{pmatrix} b = 0 \tag{A.8}$$

If n images of the model plane are observed, by stacking all the equations as (A.8), we obtain a system of $2n$ equations:

$$Gb = 0 \tag{A.9}$$

where G is a $2n$x6 matrix.

If $n \geq 3$, we will generally[1] have a unique solution b (as we have 5 intrinsic parameters to determine—f_x, f_y, $X_{pix,0}$, $Y_{pix,0}$ and γ). Once b is estimated, we can compute the camera intrinsic matrix A by (A.5).

A.5 Dealing with Radial Distortion

Up to now, we have not considered the lens radial distortion of the camera. Let (x_{pix}, y_{pix}) be the ideal (nonobservable distortion-free) pixel image coordinates, and $(\tilde{x}_{pix}, \tilde{y}_{pix})$ the corresponding real observed image coordinates. By definition of the radial distortion (2.6) and equations (2.3) and (2.8), we have:

$$\begin{array}{ccccc} \tilde{x}_{pix} & = & x_{pix} & + & (x_{pix} - X_{pix,0}) \quad d \\ \tilde{y}_{pix} & = & y_{pix} & + & (y_{pix} - Y_{pix,0}) \quad d \end{array} \tag{A.10}$$

that can be rewritten as:

$$\begin{pmatrix} x^*_{pix}\left(x^2+y^2\right) & x^*_{pix}\left(x^2+y^2\right)^2 \\ y^*_{pix}\left(x^2+y^2\right) & y^*_{pix}\left(x^2+y^2\right)^2 \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \end{pmatrix} = \begin{pmatrix} \tilde{x}_{pix} - x_{pix} \\ \tilde{y}_{pix} - y_{pix} \end{pmatrix} \tag{A.11}$$

Given m points belonging to the model plane, in n images, we can stack all the equations together to obtain a system of $2mn$ equations, that we can write in the matrix form:

$$Lk = l \tag{A.12}$$

The linear least-squares solution is given by

$$k = \left(L^T L\right)^{-1} L^T l \tag{A.13}$$

Once k_1 and k_2 are estimated, one can refine the estimate of the other five intrinsic parameters by solving (A.9), and alternate these two procedures until convergence.

[1]If among the n images, less than three have independent rotation matrices, there are not enough constraints on the intrinsic parameters provided to find an unique solution.

Another method proposed in [11] to solve this problem is what the author calls the *maximum likelihood estimation*, more complex but giving better results in less time. As it is more based on numerical methods than on 'intuition', it is not detailed here.