Formulae for camera calibration for ATTEST

Maarten Vergauwen, Frank Verbiest
K.U.Leuven ESAT-PSI-VISICS
Kasteelpark Arenberg 10
B-3001 Leuven

October 17, 2002

1 Introduction

In the ATTEST project video sequences are processed to calculate depth values for the pixels of the frames. In order to do so, the calibration of the camera has to be computed for each frame. Both intrinsic (internal) and extrinsic (external or mechanical) calibration of the camera is computed.

This document proposes a camera model to be used in the project. The camera model described in this document has been used at VISICS with success.

2 Camera model

2.1 Pin-hole model

The most used camera model in literature is the well-known **pin-hole** model. It describes the image formation process, using 3 matrices K, R and t. These matrices describe the internal parameters, the rotation and the translation of the camera respectively. A 3D point M is projected onto a 2D point m according to equation 1

$$m \simeq K[R^T| - R^T t]M \tag{1}$$

with m and M in homogeneous coordinates,

$$m = \begin{pmatrix} x \\ y \\ w \end{pmatrix} \quad M = \begin{pmatrix} X \\ Y \\ Z \\ W \end{pmatrix}$$

R and t a 3x3 rotational matrix and a 3x1 translational vector, denoting the rotation and position of the camera and K (a.k.a. the calibration matrix) the matrix with the internal parameters.

$$K = \begin{pmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix}$$

In this matrix the symbols represent the following.

- f_x and f_y are the focal length, expressed in pixels. Remark that their ratio corresponds to the aspect ratio.
- s is the skew of the pixels, a number which indicates how rectangular the pixels are. In CCDs the pixels have right angles and so the skew is zero for these cameras.
- c_x and c_y are the coordinates of the principle point. This is the intersection of the optical axis with the image plane.

Note: the symbol $\simeq$ denotes an equality up to a non-zero scale factor.

2.2 Radial distortion

The pin-hole model has the big advantage of being linear and thus very simple. One can compute the projection of a point in the image, using a simple matrix multiplication of a 4x1 vector with a 3x4 matrix.

Unfortunately, the pin-hole model does not take into account any non-linear deformations of the lens. The most important of these is the well known **radial distortion**. This distortion results in the typical effect of straight lines being projected as curves, especially at the edges of the image. The effect is non-linear because it is dependent on the distance of the pixel to the centre of the distortion in a quadratic (or higher order) way.

The camera calibration algorithms, developed by VISICS estimate parameters to model this radial distortion and experience has shown that much better results can be achieved if this distortion is taken into account.

The simplest way of dealing with radial distortion, is to apply an extra term in the projection equation, left from the internal calibration matrix K. Following the image formation process, however, it is more natural to put the distortion parameters in between the mechanical (rotational and translational) and internal part. Equation 2 shows the general formula.

$$m \simeq K \ \mathbf{d} \ \left([R^T| - R^T t]M\right) \tag{2}$$

Equation 2 can be read from right to left and shows that the 3D point M is first put into the coordinate system of the camera by multiplication with the 3x4 matrix with the extrinsic calibration. Then the radial distorion is applied after which the intrinsic calibration matrix puts the point on the image.

Let's define

$$U = \begin{pmatrix} U_x \\ U_y \\ U_w \end{pmatrix} = [R^T| - R^T t] \begin{pmatrix} X \\ Y \\ Z \\ W \end{pmatrix}$$

The distance from the optical axis r is then

$$r^2 = \left(\frac{U_x}{U_w}\right)^2 + \left(\frac{U_y}{U_w}\right)^2$$

If the radial distortion is modeled, using three or more parameters κ_1, κ_2, κ_3, $\ldots$, then we define D as

$$D = \begin{pmatrix} D_x \\ D_y \\ D_w \end{pmatrix} \simeq \begin{pmatrix} (1 + \kappa_1 r^2 + \kappa_2 r^4 + \kappa_3 r^6 + \ldots)\, U_x \\ (1 + \kappa_1 r^2 + \kappa_2 r^4 + \kappa_3 r^6 + \ldots)\, U_y \\ U_w \end{pmatrix} \tag{3}$$

Then finally the projection m is

$$m \simeq KD$$

3 Backprojecting

The previous section described the process of image forming in detail. It shows how to project a 3D point into the image. The reverse process is also used in ATTEST. This process is called backprojecting. It essentially computes a 3D line in space through the camera centre for a 2D image point. All points on this line project to the same 2D point. As a consequence the 3D point we are looking for also lies on this line. The exact distance on the line between the camera centre and the 3D point can be read out from the depth map.

If one uses the simple pin-hole model, constructing this 3D line is an easy problem. The use of radial distortion parameters makes our life more difficult, though.

Backprojecting can be seen as reading equation 2 from left to right and inverting every step. The first term one encounters is the internal calibration matrix K. This step is easily inverted. Let us define D as

$$D = \begin{pmatrix} D_x \\ D_y \\ D_w \end{pmatrix} \simeq K^{-1} \begin{pmatrix} x \\ y \\ w \end{pmatrix} = K^{-1} m$$

The next step consists of undoing the radial distortion. We want to find the point U which is the undistorted point of D. We know that D comes from U using equation 3. One can solve this equation from right to left to compute U from D.

Now we have the undistorted coordinates, we can easily compute the backprojected line of these coordinates, using the external calibration matrix $[R^T| - R^T t]$.

If many backprojections are necessary, it might be interesting to not always invert equation 3 but to build up a lookup table.

4 Calibration file

In this section we propose a camera file which contains all necessary parameters for the camera model described in this text.

The format of the file is as follows.

- The 3x3 matrix K (internal calibration of the camera)
- 3 radial distortion parameters (κ_1, κ_2, κ_3). We have experienced that two parameters are normally more than enough, so the third one will probably be 0.
- The 3x3 matrix R (rotation of the camera)
- The vector of length 3 t (translation of the camera)

This is an example of the format

```
900.9061005324038 0 384.6688757203483
0 922.2912273093366 284.3854530460084
0 0 1

-0.2342116924069423 0.2410398499662985 0

-0.9997301800819204 -0.006478148084638816 -0.02230696372804397
-0.006613193829926063 0.9999602187356657 0.005985533684412003
0.0222673011552502 0.006131438943094297 -0.9997332508003063

0.4970740570694551 -0.115762613082715 0.9820705504640316
```