# COL 780 - Assignment 3

## Introduction:

A 2D point is denoted by m = [u, v] T . A 3D point is denoted by M = [X, Y, Z] T . We use x to denote the augmented vector by adding 1 as the last element: me = [u, v, 1]T and M = [X, Y, Z, 1]T . A camera is modeled by the usual pinhole: the relationship between a 3D point M and its image projection m is given by

$$s\tilde{\mathbf{m}} = \mathbf{A}\begin{bmatrix}\mathbf{R} & \mathbf{t}\end{bmatrix}\tilde{\mathtt{M}}\,, \tag{1}$$

where s is an arbitrary scale factor, (R, t), called the extrinsic parameters, is the rotation and translation which relates the world coordinate system to the camera coordinate system, and A, called the camera intrinsic matrix, is given by

$$\mathbf{A} = \begin{bmatrix}\alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1\end{bmatrix}$$

with (u0, v0) the coordinates of the principal point, α and β the scale factors in image u and v axes, and γ the parameter describing the skewness of the two image axes.

Given an image of the model plane, an homography can be estimated. Let's denote it by H = [ h1 h2 h3 ] . From (2), we have

$$\begin{bmatrix}\mathbf{h}_1 & \mathbf{h}_2 & \mathbf{h}_3\end{bmatrix} = \lambda\mathbf{A}\begin{bmatrix}\mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t}\end{bmatrix}\,,$$

where λ is an arbitrary scalar. Using the knowledge that r1 and r2 are orthonormal, we have

$$\mathbf{h}_1^T\mathbf{A}^{-T}\mathbf{A}^{-1}\mathbf{h}_2 = 0 \tag{3}$$

$$\mathbf{h}_1^T\mathbf{A}^{-T}\mathbf{A}^{-1}\mathbf{h}_1 = \mathbf{h}_2^T\mathbf{A}^{-T}\mathbf{A}^{-1}\mathbf{h}_2\,. \tag{4}$$

These are the two basic constraints on the intrinsic parameters, given one homography. Because a homography has 8 degrees of freedom and there are 6 extrinsic parameters (3 for rotation and 3 for translation), we can only obtain 2 constraints on the intrinsic parameters.

**Camera Calibration:**

We start with an analytical solution, followed by a nonlinear optimization technique based on the maximum likelihood criterion. Finally, we take into account lens distortion, giving both analytical and nonlinear solutions.

Let,

$$\mathbf{B} = \mathbf{A}^{-T}\mathbf{A}^{-1} \equiv \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{12} & B_{22} & B_{23} \\ B_{13} & B_{23} & B_{33} \end{bmatrix}$$
$$= \begin{bmatrix} \frac{1}{\alpha^2} & -\frac{\gamma}{\alpha^2\beta} & \frac{v_0\gamma-u_0\beta}{\alpha^2\beta} \\ -\frac{\gamma}{\alpha^2\beta} & \frac{\gamma^2}{\alpha^2\beta^2}+\frac{1}{\beta^2} & -\frac{\gamma(v_0\gamma-u_0\beta)}{\alpha^2\beta^2}-\frac{v_0}{\beta^2} \\ \frac{v_0\gamma-u_0\beta}{\alpha^2\beta} & -\frac{\gamma(v_0\gamma-u_0\beta)}{\alpha^2\beta^2}-\frac{v_0}{\beta^2} & \frac{(v_0\gamma-u_0\beta)^2}{\alpha^2\beta^2}+\frac{v_0^2}{\beta^2}+1 \end{bmatrix} . \tag{5}$$

Here, B is symmetric, defined by a 6D vector.

$$\mathbf{b} = [B_{11}, B_{12}, B_{22}, B_{13}, B_{23}, B_{33}]^T . \tag{6}$$

Let the i th column vector of H be hi = [hi1, hi2, hi3] T . Then, we have

$$\mathbf{h}_i^T\mathbf{B}\mathbf{h}_j = \mathbf{v}_{ij}^T\mathbf{b} \tag{7}$$

$$\mathbf{v}_{ij} = [h_{i1}h_{j1}, h_{i1}h_{j2} + h_{i2}h_{j1}, h_{i2}h_{j2}, \\ h_{i3}h_{j1} + h_{i1}h_{j3}, h_{i3}h_{j2} + h_{i2}h_{j3}, h_{i3}h_{j3}]^T .$$

Therefore, the two fundamental constraints (3) and (4), from a given homography, can be rewritten as 2 homogeneous equations in b:

$$\begin{bmatrix} \mathbf{v}_{12}^T \\ (\mathbf{v}_{11} - \mathbf{v}_{22})^T \end{bmatrix} \mathbf{b} = \mathbf{0} \,. \tag{8}$$

If n images of the model plane are observed, by stacking n such equations as (8) we have,

$$Vb=0 \tag{9}$$

where V is a 2n×6 matrix.

Once b is estimated, we can compute all camera intrinsic matrix A.

B = λA−T A with λ an arbitrary scale. Without difficulty , we can uniquely extract the intrinsic parameters from matrix B.

$$\begin{aligned}
v_0 &= (B_{12}B_{13} - B_{11}B_{23})/(B_{11}B_{22} - B_{12}^2) \\
\lambda &= B_{33} - [B_{13}^2 + v_0(B_{12}B_{13} - B_{11}B_{23})]/B_{11} \\
\alpha &= \sqrt{\lambda/B_{11}} \\
\beta &= \sqrt{\lambda B_{11}/(B_{11}B_{22} - B_{12}^2)} \\
\gamma &= -B_{12}\alpha^2\beta/\lambda \\
u_0 &= \gamma v_0/\beta - B_{13}\alpha^2/\lambda \,.
\end{aligned}$$

Once A is known, the extrinsic parameters for each image is  readily computed.

$$\begin{aligned}\mathbf{r}_1 &= \lambda \mathbf{A}^{-1}\mathbf{h}_1 \\ \mathbf{r}_2 &= \lambda \mathbf{A}^{-1}\mathbf{h}_2 \\ \mathbf{r}_3 &= \mathbf{r}_1 \times \mathbf{r}_2 \\ \mathbf{t} &= \lambda \mathbf{A}^{-1}\mathbf{h}_3\end{aligned}$$

with,

$$\lambda = 1/\|\mathbf{A}^{-1}\mathbf{h}_1\| = 1/\|\mathbf{A}^{-1}\mathbf{h}_2\|.$$

**Distortion Coefficients:**

Let (u, v) be the ideal (non observable distortion-free) pixel image coordinates, and (˘u, v˘) the corresponding real observed image coordinates. The ideal points are the projection of the model points according to the pinhole model. Similarly, (x, y) and (˘x, y˘) are the ideal (distortion-free) and real (distorted) normalized image coordinates.

$$\begin{aligned}\breve{x} &= x + x[k_1(x^2+y^2) + k_2(x^2+y^2)^2] \\ \breve{y} &= y + y[k_1(x^2+y^2) + k_2(x^2+y^2)^2]\ ,\end{aligned} \qquad (10,11)$$

where k1 and k2 are the coefficients of the radial distortion. The center of the radial distortion is the same as the principal point.

Now,

$$\begin{bmatrix}(u-u_0)(x^2+y^2) & (u-u_0)(x^2+y^2)^2 \\ (v-v_0)(x^2+y^2) & (v-v_0)(x^2+y^2)^2\end{bmatrix}\begin{bmatrix}k_1 \\ k_2\end{bmatrix} = \begin{bmatrix}\breve{u}-u \\ \breve{v}-v\end{bmatrix}.$$

As , u˘ = u0 + αx˘ + γy˘ and v˘ = v0 + βy˘ and assuming γ = 0,
Now , Given m points in n images, we can stack all equations together to obtain in total 2mn equations, or in matrix form as Dk = d, where k = [k1, k2] T . The linear least-squares solution is given by:

$$\mathbf{k} = (\mathbf{D}^T\mathbf{D})^{-1}\mathbf{D}^T\mathbf{d}\ . \qquad (12)$$

**REFINING:**

Now , refine the estimate of the other parameters by solving the equation

$$\sum_{i=1}^{n}\sum_{j=1}^{m}\|\mathbf{m}_{ij} - \breve{\mathbf{m}}(\mathbf{A}, k_1, k_2, \mathbf{R}_i, \mathbf{t}_i, \mathtt{M}_j)\|^2\ , \qquad (13)$$

where m˘ (A, k1, k2, Ri , ti , Mj ) is the projection of point Mj in image i according to equation (2), followed by distortion according to (11) and (12).
An initial guess of k1 and k2 can be obtained simply by setting them to 0.

Thus in summary ,

- Print a pattern and attach it to a planar surface.
- Take a few images of the model plane under different orientations by moving either the plane or the camera.
- Detect the feature points in the images.
- Estimate the five intrinsic parameters and all the extrinsic parameters using the closed-form solution.
- Estimate the coefficients of the radial distortion by solving the linear least-squares (12).
- Refine all parameters by minimizing (13)

Now , we need to set up a 3D coordinate  system on a 2D image plane.We do this by :

- Defining an objp2 array of object points in 3D space
- Using the rotation and translational matrix i.e the extrinsic parameters.for the detected corners (corners2) in the undistorted image.
- Finally, we call the cv.projectPoints() function to project the 3D points onto the 2D image plane (imgpts2 and imgpts3).
- We then define a draw function, draw2() , that uses the detected corners and projected 3D points to draw a coordinate system and a cube in the image, respectively.
- Return: The undistorted image with the drawn coordinate system and cube is returned at the end.

**OUTPUT IMAGES LINK**—-  output_assign3