TP6: Projective Geometry and camera calibration

Loic Simon Chaohui Wang Panagiotis Koutsourakis

28 May 2008

1 Theory

1.1 Projective Space

The projective space of dimension n, $\mathbb{P}^n$ is the quotient[1] space of $\mathbb{R}^{n+1} \setminus \{0_{n+1}\}$ defined by the following equivalence relation:

$$[x_1, \ldots, x_{n+1}]^t \sim [x'_1, \ldots, x'_{n+1}]^t \Leftrightarrow \exists \lambda \neq 0, [x_1, \ldots, x_{n+1}]^t = \lambda [x'_1, \ldots, x'_{n+1}]^t$$

Note that this relation is clearly reflexive, symmetric and transitive.

The points of $\mathbb{P}^n$ that satisfy $x_{n+1} \neq 0$ have an equivalent in the euclidian space $\mathbb{R}^n$ $[\frac{x_1}{x_{n+1}}, \ldots \frac{x_n}{x_{n+1}}]^t$. The points that satisfy $x_{n+1} = 0$ don't have and Euclidean equivalent and are called *the points at infinity*.

1.2 The pinhole camera model

The pinhole camera model allows us to describe the process of the acquisition of an image by the projection of the 3D points, to some 2D points situated on the retinal plane. Let C be the optical center of the camera. The projection of a 3D point M is the intersection of the optic ray CM with the retinal plane as you can see in figure 1.

Let $M = [x\ y\ z]^t$ be a point of the Euclidean 3D space and $m = [u\ v]^t$ its projection. Let $\widetilde{M} = [x\ y\ z\ 1]^t$ and $\widetilde{m} = [u\ v\ 1]$ be the corresponding points in homogeneous coordinates. In order to formalize the projection process we write:

$$\mathcal{P}\widetilde{M} = \widetilde{m}$$

where $\mathcal{P}$ is a 3×4 projection matrix.

[1] The quotient space is the set of equivalence classes.

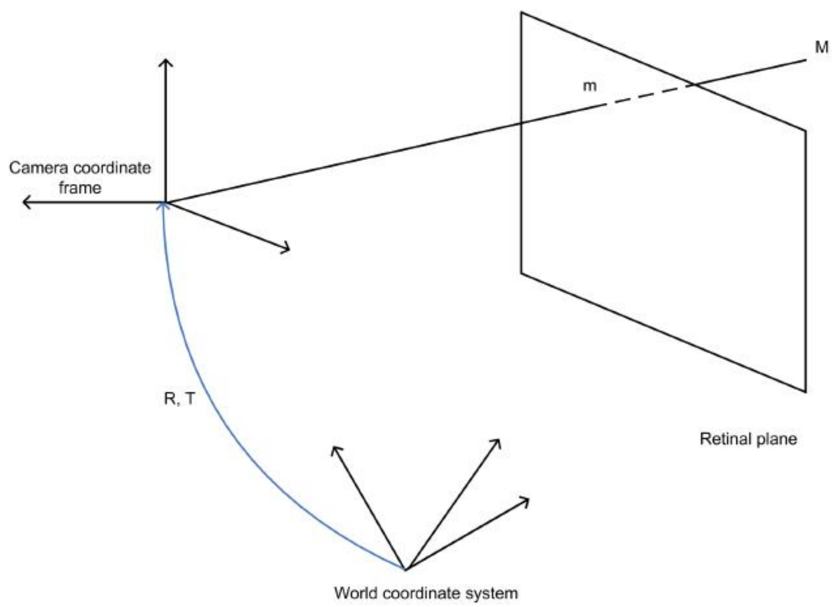

Figure 1: The pinhole camera model

1.3 Decomposition of the projection matrix

In this part we define 2 coordinate frames: one linked to the scene (which we can choose) and the other linked to the camera. The origin of the camera reference frame is the center of the camera and its axes are the axes of the retinal plane and the optical axis (this is the axis that passes from the camera center and is normal to the retinal plane).

It can be shown that the projection matrix $\mathcal{P}$ is decomposed as follows:

$$\mathcal{P} = \mathbf{K}\left[\mathbf{R}|\mathbf{T}\right]$$

where:

- $\mathbf{K}$ is a 3×3 matrix that contains the *intrinsic* parameters of the camera (these are the parameters that depend only on the internal configuration of the camera). This matrix describes the reference frame of the retinal plane. We can write it in the following form:

$$\mathbf{K} = \begin{bmatrix} f_u & \gamma & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

  f_u and f_v are the focal distances expressed in pixels. $[u_0\ v_0]$ are the coordinates of the principal point, i.e. the intersection point of the retinal plane with the optical axis. Finally γ is called *skew* parameter and is 0 for most normal cameras.

- $[\mathbf{R}|\mathbf{T}]$ represents the pose of the camera in the scene reference frame. This rigid transformation allows us to transform the 3D points of the scene reference frame to the camera reference frame. $\mathbf{R}$ is a 3×3 rotation matrix and $\mathbf{T}$ is a 3×1 translation vector. $\mathbf{R}$ and $\mathbf{T}$ are the *extrinsic* parameters of the camera.

Note that the projection matrix depends on 11 parameters: 5 intrinsic parameters and 6 extrinsic parameters (3 for rotation and 3 for translation). The process of *calibration* of a camera consists in estimating its intrinsic and/or extrinsic parameters.

1.4 Calculation of the projection matrix from 2D-3D correspondences

Returning to the following equation:

$$\mathcal{P}\widetilde{M} = \widetilde{m}$$

we can write:

$$\mathcal{P} = \begin{bmatrix} \mathcal{P}_1^t \\ \mathcal{P}_2^t \\ \mathcal{P}_3^t \end{bmatrix}$$

That means that $\mathcal{P}_i$ is the ith line of $\mathcal{P}$, written as column vector and therefore is of dimension 4×1. For example $\mathcal{P}_1 = [\mathcal{P}_{11}\ \mathcal{P}_{12}\ \mathcal{P}_{13}\ \mathcal{P}_{14}]^t$. It is important to understand that this equation gives two independent equations depending on the elements of $\mathcal{P}$:

$$\begin{aligned} \frac{\langle \mathcal{P}_1, \widetilde{M} \rangle}{\langle \mathcal{P}_3, \widetilde{M} \rangle} &= u \\ \frac{\langle \mathcal{P}_2, \widetilde{M} \rangle}{\langle \mathcal{P}_3, \widetilde{M} \rangle} &= v \end{aligned} \tag{1}$$

where (u, v) are the euclidean coordinates of the point $\widetilde{m}$.

Equations (1) can be written in the following form:

$$\begin{aligned} \langle \mathcal{P}_1, \widetilde{M} \rangle - u \langle \mathcal{P}_3, \widetilde{M} \rangle &= 0 \\ \langle \mathcal{P}_2, \widetilde{M} \rangle - v \langle \mathcal{P}_3, \widetilde{M} \rangle &= 0 \end{aligned} \tag{2}$$

Equations (2) are linear with respect to the elements of $\mathcal{P}$. If there are n correspondences, we obtain a homogeneous linear system $\mathbf{AP} = 0$, where $\mathbf{A}$ is one $2n \times 12$ matrix and

$$\mathbf{P} = [\mathcal{P}_{11}\ \mathcal{P}_{12}\ \mathcal{P}_{13}\ \mathcal{P}_{14}\ \mathcal{P}_{21}\ \mathcal{P}_{22}\ \mathcal{P}_{23}\ \mathcal{P}_{24}\ \mathcal{P}_{31}\ \mathcal{P}_{32}\ \mathcal{P}_{33}\ \mathcal{P}_{34}]^t$$

In order to avoid the trivial solution $\mathbf{P} = 0$ and knowing that the projection matrix can only be determined up to a multiplying factor, we are going to reformulate the problem of solving the linear system in the form of an optimization problem:

$$\min_{\|\mathbf{P}\|^2=1} \|\mathbf{AP}\|^2 \tag{3}$$

We can show that the solution to this problem can be obtained by computing the singular value decomposition (SVD) of the matrix $\mathbf{A}$ and taking the vector corresponding to the smallest singular value.

1.5 Extraction of intrinsic and extrinsic parameters from the projection matrix

We remind that the projection matrix can be decomposed as follows:

$$\mathcal{P} = \mathbf{K}\,[\mathbf{R}|\mathbf{T}]$$

The objective is to find $\mathbf{K}$, $\mathbf{R}$ and $\mathbf{T}$ from $\mathcal{P}$. To this end we are going to use the QR decomposition. We remind briefly here that we can decompose one square matrix M as $M = QR$ where Q is an orthogonal matrix and R is an upper triangular matrix (do not comfuse this R with the rotation matrix $\mathbf{R}$).

Let A be the upper left 3×3 submatrix of $\mathcal{P}$, i.e. it contains the elements $\mathcal{P}_{ij}$, $1 \leq i \leq 3$, $1 \leq j \leq 3$. Thus $A = \mathbf{KR}$. We perform a QR decomposition on A^{-1}:

$$A^{-1} = QL$$

with Q orthogonal and L upper triangular. This gives:

$$A = L^{-1}Q^{-1} = L^{-1}Q^{t}$$

where $\mathbf{K} = L^{-1}$ and $\mathbf{R} = Q^t$. When computing $\mathbf{K}$ and $\mathbf{R}$ don't forget to normalize $\mathbf{K}$ (and thus $\mathcal{P}$) in order to have $K_{33} = 1$. After that, finding the translation vector is easy.

2 Computation

Take care: an $m \times n$ mathematical matrix, is represented in the CImg library by an $n \times m$ `CImg<double>` *object.*

2.1 Computation of the projection matrix

Complete the function `computeProjectionMatrix` that takes as arguments a set of 3D points in projective coordinates `points3D` (of dimension $4 \times n$) and the set of corresponding 2D points of the image in Euclidean coordinates `points2D` (of dimensions $2 \times n$). This function calculates the matrix `projectionMatrix`.

2.2 Extraction of intrinsic and extrinsic parameters

Complete the function `computeParameters` that takes as arguments the projection matrix `projectionMatrix` and calculates the intrinsic parameters matrix `intrinsicParameters` ($\mathbf{K}$) the rotation matrix `rotation` ($\mathbf{R}$) and the translation vector `translation` ($\mathbf{T}$). This function should call the auxiliary function `QRdecomposition` that is already coded, that performs the QR decomposition of a matrix.