Pose estimation

PIV, 2014/15

This note addresses the estimation of camera pose (R, t) with respect to the world coordinate system, assuming that we know the projection of a set of marks extracted from an object with known geometry (calibration object). Two cases are considered: i) non coplanar marks and ii) coplanar marks.

The inner parameters of the camera, K, are assumed to be known.

1 Pose estimation from non coplanar points

Let us assume we know N non coplanar points $p^1, \dots, p^N$, and their projections on the image plane $x^1, \dots, x^N$. Since the intrinsic parameters of the camera are known, we can normalize the projected points according to

$$x^i \leftarrow K^{-1} x^i$$

and estimate the camera matrix $\hat{P}$ e.g., using the DLT method.

Since the points were corrected by K^{-1}, the camera matrix P is given by

$$P = [R \mid t] ,$$

where R, t are the pose parameters. If we split the estimated camera matrix $\hat{P}$ into two blocks of sizes 3×3, 3×1

$$\hat{P} = [Q \mid q] ,$$

the two matrices $P, \hat{P}$ should be approximately equal, up to a scale factor. Therefore,

$$R \approx \alpha Q ,$$

$$t \approx \alpha q ,$$

where α can be obtained by computing the determinant of R and Q, leading to

$$\alpha = \frac{1}{det(Q)^{1/3}} .$$

Unfortunately, the matrix R obtained by the expression $R = \alpha Q$ is not a rotation matrix, *i.e.*, it does not have the required properties

$$R^T R = R R^T = I \qquad \det(R) = 1$$

We can therefore ask what is the rotation matrix R nearest to αQ, according to a suitable metric. A convenient metric for this problem is the Frobenious norm

$$\|Q - R\|_F = \sqrt{\sum_{i,j} (q_{ij} - r_{ij})^2}$$

The optimization problem formulated above is a special case of the Procrustes problem solved in the sixties which we will briefly summarise.

Procrustes problem

We wish to solve the following optimization problem

$$R = \arg\min_{\Omega} \|A\Omega - B\|_F \ , \quad \text{subject to } \Omega^T\Omega = I \ , \quad \det \Omega = 1 \ .$$

The solution is based on the singular value decomposition of $A^T B$: $A^T B = UDV^T$. The rotation matrix that solves the Procrustes problem is given by

$$R = U \Sigma V^T$$

where Σ is a diagonal matrix of ones, except the last entry which can take the value $+1$ or -1, depending on the sign of $\det(UV^T)$

$$\Sigma = \begin{bmatrix} 1 & 0 & ... & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & sign\left(\det(UV^T)\right) \end{bmatrix} .$$

The pose estimation problem formulated above, is a special case of the Procrustes problem with $A = I, B = Q$. Therefore, we have to compute the SVD decomposition of Q and apply the above provedure: $R = U\Sigma V^T$.

2 Pose estimation from coplanar points

If the marks are coplanar (*e.g.*, they are all located on a wall) the above procedure cannot be used since we do not have enough information to estimate the camera matrix P.

To simplify the problem, we will assume that the world coordinate system is carefully chosen such that, the plane associated to the marks can be described by $p_z = 0$. In this case, the camera model becomes

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ 0 \\ 1 \end{bmatrix} .$$

The third column of P multiplies by zeros. Therefore, it can be removed,

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{14} \\ p_{21} & p_{22} & p_{24} \\ p_{31} & p_{32} & p_{34} \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ 1 \end{bmatrix}$$

and we obtain an homography between the plane coordines (p_x, p_y) and the image coordinates leading to, $\lambda \tilde{x} = H\tilde{X}$. Matrix H can be estimated from the experimental data by linear or non linear methods.

Since, $P = [R \mid t]$, and H is obtained from P by removing the third column, we conclude that

$$\alpha h_{c1} = r_{c1} \qquad \alpha h_{c2} = r_{c2} \qquad \alpha h_{c3} = t$$

where h_{ci} and r_{ci} are the ith columns of H and R, respectively.

The columns of matrix R should have unit norm and should be orthogonal. These constraints can be met by a set of operations:

i) normalize the first column of H,

$$r_{c1} = \mathcal{N}(h_{c1}) \ ,$$

ii) orthogonalize h_{c2} with respect to r_{c1},

$$r_{c2} = \mathcal{N}(h_{c2} - (h_{c2} \cdot r_{c1})\ r_{c1}) \ ,$$

iii) obtain the third column bythe external product of the two first columns

$$r_{c2} = r_{c2} \times r_{c2} \ .$$

In these expressions, $\mathcal{N}(v) = v/\|v\|$, denotes the normalization of the input vector v in order to obtain a vector with the same direction and norm 1.