P3.5P: Pose Estimation with Unknown Focal Length

Changchang Wu
Google Inc.

The problem of pose estimation with a single unknown focal length is a widely adopted approximation in the reconstruction of uncalibrated photos (e.g. Internet photos). This problem has 7 degrees of freedom, thus requires a minimum of 3.5 image points of 4 known 3D points to recover the unknowns, which this paper refers to as **P3.5P**. Unfortunately, the existing methods require 4 full image points to solve the camera pose and focal length [1, 2, 3]. This paper presents the first general solution to the P3.5P problem with up to 10 solutions.

Existing methods for camera parametrization do not allow to solve the general P3.5P problem efficiently. In particular, when focal length f is part of camera parametrization, there exists a two-fold redundancy with negative focal lengths. By decomposing camera rotation R into

$$R = R_\theta R_\rho = R(z,\,\theta)\;R(\Phi,\,\rho) \tag{1}$$

where R_θ is a rotation by θ around the z-axis and R_ρ is a rotation by ρ around a unit axis Φ in the xy plane ($\Phi \perp z$), this paper introduces a new compact camera parametrization

$$P = \left[K_\theta\, R_\rho \;\middle|\; T\right],$$

where

$$K_\theta = \begin{bmatrix} f_c & -f_s & \\ f_s & f_c & \\ & & 1 \end{bmatrix} = \begin{bmatrix} f\cos\theta & -f\sin\theta & \\ f\sin\theta & f\cos\theta & \\ & & 1 \end{bmatrix}. \tag{2}$$

By using a quaternion vector $[1,\,q_x,\,q_y,\,0]^\top$ to represent the rotation R_ρ, the $P3.5P$ problem has the following 7 unknowns to solve

$$\{f_c,\,f_s,\,q_x,\,q_y,\,t_x,\,t_y,\,t_z\} \tag{3}$$

This new parametrization successfully eliminates the two-fold redundancy of negated focal lengths by merging two alternative solutions : $f_c = f\cos\theta = (-f)\cos(\theta+\pi)$ and $f_s = f\sin\theta = (-f)\sin(\theta+\pi)$.

By applying the new parametrization into the perspective projection, it is then possible to eliminate the translation vector T. Let R^1_ρ, R^2_ρ and R^3_ρ be the three rows of R_ρ, we can derive a simple constraint between any point (x_i, y_i) and two other coordinate x_j and y_k:

$$\begin{aligned} 0 = {} & (y_i - y_k)(f_c R^1_\rho - f_s R^2_\rho - x_j R^3_\rho)(X_j - X_i) - \\ & (x_i - x_j)(f_s R^1_\rho + f_c R^2_\rho - y_k R^3_\rho)(X_k - X_i), \end{aligned} \tag{4}$$

which is of degree 3 in $\{f_c,\,f_s,\,q_x,\,q_y\}$. Since the polynomial equation encodes the geometric relationship between **4** image coordinates, it is called the **quadruple constraint** by $\{(x_i,y_i),x_j,y_k\}$.

There are exactly 4 linearly independent quadruple constraints. Because f_c and f_s appear linearly in these constraints, we can rewrite the 4 quadruple constraints as the multiplication of a 4×3 degree 2 polynomial matrix $F(q_x,q_y)$ with vector $[f_c,\,f_s,\,1]^\top$.

$$\underbrace{F(q_x,q_y)}_{4\times 3} \begin{bmatrix} f_c \\ f_s \\ 1 \end{bmatrix} = 0 \tag{5}$$

Because these equations have non-trivial solutions, $F(q_x,q_y)$ must be rank deficient, and each 3×3 sub-matrix of $F(q_x,q_y)$ is has a determinant of 0. By computing the determinants of the 4 possible 3×3 sub-matrices of $F(q_x,q_y)$, we are able to eliminate f_c and f_s and derive 4 polynomials in just two unknowns $\{q_x,q_y\}$ of degree 6. In comparison, the polynomials used in recent general P4P methods [1] and [3] are of degree 8 and 7 respectively.

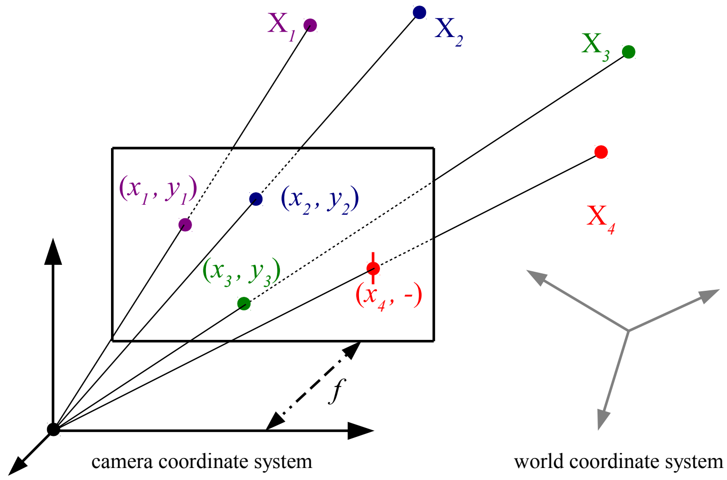

Figure 1: The P3.5P problem. The input are the 3 image points (x_1,y_1) (x_2,y_2) (x_3,y_3) plus either x_4 or y_4 of the 4-th image point and their corresponding 3D points $\{X_1,X_2,X_3,X_4\}$ in the world coordinate system. The 7 image coordinates give sufficient constraints for this paper to recover the unknown focal length and camera pose.

Standard Gröbner basis eigenvalue method is used in this paper to solve the polynomials of $\{q_x,q_y\}$. The lowered polynomial degrees lead to a reduced elimination template size of 20×30, which is much smaller than other P4P methods and yields an improved speed accordingly. See Table 1 for the details. For each solution of q_x and q_y, the remaining unknowns can be straight-forwardly solved using singular value decomposition. Experiments show the proposed P3.5P has similar accuracy compared to existing P4P methods [1, 3], and it works well in real reconstruction systems.

Now the P3.5P algorithm has used 7 image coordinates to find up to 10 possible solutions, and there is still one remaining image coordinate to use. We can verify the recovered solutions and filter out invalid ones using the remaining coordinate. Normally only a single solution remains after the filtering for valid 4 image points, and no solution exist for outliers. This simple filtering step allows for RANSAC to test fewer solutions and leads to another speedup for real applications.

Solver	Polynomial solving method	Time
Our P3.5P	GB, 20×30 G-J elimination	0.109ms
Ratio [1]	GB, 53×63 G-J elimination	0.336ms
	GB, 86×96 G-J elimination	0.929ms
	GB, 139×153 G-J elimination	3.320ms
Zheng [3]	GB, 36×52 G-J elimination	0.257ms
	polyeig	1.648ms
	Characteristic polynomial	0.067ms

Table 1: Comparison of the solver speed (the polynomial solving step). The timings for the GB-based solvers are benchmarked using G-J elimination and Eigen decomposition of random matrices.

[1] Martin Bujnák. Algebraic solutions to absolute pose problems, 2012.

[2] Bill Triggs. Camera pose and calibration from 4 or 5 known 3d points. In *CVPR*, 1999.

[3] Yinqiang Zheng, Shigeki Sugimoto, Imari Sato, and Masatoshi Okutomi. A general and simple method for camera pose and focal length determination. In *CVPR*, 2014.

This is an extended abstract. The full paper is available at the Computer Vision Foundation.