# A Note on the Difference between the Camera Resection and the PnP Problem[1)]

SUN Feng-Mei

(*College of Science, North China University of Technology, Beijing* 100041)
(E-mail: fmsun@ncut.edu.cn)

**Abstract** The camera resection and the PnP problem are both of great importance in computer vision field. However, these two fundamentally different problems are often mistaken one another in the literature. In this short note, the essential differences of these two problems are clarified, and such a clarification seems helpful for those who are working on the problems.

**Key words** PnP problem, camera calibration

## 1 Introduction

The PnP problem (Perspective-N-Points problem) is called "space resection" in photogrammetry community. That is why people often mistake it as camera resection in computer vision field. Moreover, people sometimes use results drawn from one problem to judge peers' works on the other, unreasonably criticize others' valuable works, and rubbish them unfairly. Hence a clarification on the differences of these two problems seems necessary and helpful.

Of course, mistaking and confusions are not accidental in literature. The worth of preferential treatment of these two problems lies primarily in their high importance in object pose determination and camera calibration in computer vision field.

## 2 The PnP problem

The PnP problem was first formally introduced by Fishler and Bolles in 1981[1], and later extensively studied by others, to cite a few[1~11]. There are two different definitions for the PnP problem in the literature[9]. One is the distance based definition, the other is the transformation based definition.

### 2.1 Distance based definition

In [1], the PnP problem was defined as:

Given the relative spatial locations of $n$ control points, and given the angles to every pairs of control points from the perspective center (the camera's optical center), find the lengths of the line segments joining the perspective center to each of the control points.

Since the distances from the perspective center to the control points should be determined in this definition, it is called "distance based definition". The distance based definition was exemplified in Harallick's work about the P3P problem in [2], which can be summarized as follows.

Given the three 3D control points $\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3$, the three distances $a = \|\boldsymbol{p}_2 - \boldsymbol{p}_3\|, b = \|\boldsymbol{p}_1 - \boldsymbol{p}_3\|, c = \|\boldsymbol{p}_1 - \boldsymbol{p}_2\|$, the corresponding three 2D image points $\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3$ of points $\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3$, and the camera's intrinsic parameters, find the unknown distances $s_1, s_2, s_3$ of the points $\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3$ from the camera's optical center.

### 2.2 Transformation based definition

In the literature, the PnP problem is also defined as to determine the 3D rigid transformation from the object-centered frame to the camera centered frame. In [3], the PnP problem was defined as:

Given a set of points with their coordinates in an object-centered frame and their corresponding projections onto an image plane and given the intrinsic camera parameters, find the transformation matrix (three rotation parameters and three translation parameters) between the object frame and the camera frame.

Since the transformation matrices should be determined in this definition, it is called "transformation based definition".

1) Supported by National Natural Science Foundation of P. R. China (60375006) and the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences (NLPR-2003-05)
Received February 5, 2004; in revised form August 16, 2004

### 2.3 The two definitions of the PnP problem are generally different

The basic difference is that the configurations of control points which all satisfy the distance based definition cannot sometimes be related each other by a 3D rotation and a translation. As shown for the P4P problem in Fig. 1, both configurations $(A,B,C,D)$ and $(A,B,C,D')$ satisfy the distance based definition, where $O$ is the optical center, $a,b,c,d$ are the 4 image points, line $OD$ is perpendicular to plane $ABC$, $E$ is the intersecting point, and $|DE| = |D'E|$. In other words, both $(|OA|,|OB|,|OC|,|OD|)$ and $(|OA|,|OB|,|OC|,|OD'|)$ are positive solutions to the P4P problem under the distance based definition. However, it is impossible to transform configuration $(A,B,C,D)$ into configuration $(A,B,C,D')$ by a rotation and a translation since configuration $(A,B,C,D)$ is the reflection of configuration $(A,B,C,D')$ with respect to plane $(ABC)$, *i.e.*, configuration $(A,B,C,D)$ and configuration $(A,B,C,D')$ cannot both be solutions to the P4P problem under the transformation based definition.

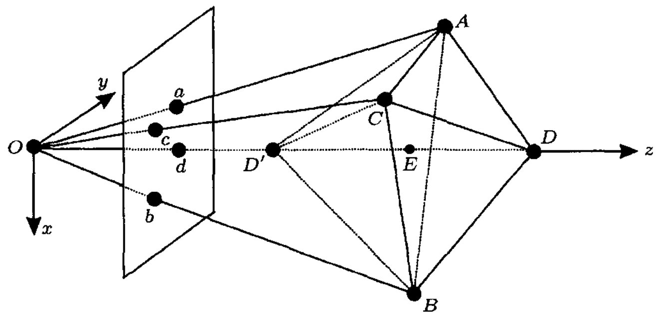


Fig. 1 Both configuration $(A,B,C,D)$ and configuration $(A,B,C,D')$ are positive solutions to the P4P problem under the distance based definition, however, they cannot both be solutions to the P4P problem under the transformation based definition

Generally speaking, if all the control points are coplanar, or the problem has a unique solution, the two definitions are equivalent. Otherwise, they are generally not equivalent. In [13], it was shown that for the P3P problem, the two definitions were equivalent, but for the PnP problem with $N > 3$, the distance-based definition was equivalent to the so-called orthogonal transformation based definition. In other words, if the rotation matrix in the transformation based definition is relaxed to an orthogonal matrix, the two definitions will be equivalent.

## 3 Camera resection

Camera resection is defined to determine the consistent projection matrices for a given set of correspondences from 3D space control points to 2D image points [12, p. 516], where the coordinates of the image points and those of the control points are both known. For example, for a given set of point correspondences $\{\boldsymbol{u}_i \Leftrightarrow \boldsymbol{X}_i \; i = 1,2,3,\cdots,N\}$, the problem of camera resection is to determine all consistent projection matrices $P_{3\times 4}$ such that all the following equations hold:

$$\lambda_i \boldsymbol{u}_i = P_{3\times 4}\boldsymbol{X}_i, \quad i = 1,2,3,\cdots N$$

where $\lambda_i$s are unknown scale factors; $\boldsymbol{u}_i$s are 2D image points in homogenous coordinates, $\boldsymbol{X}_i$s are 3D control points in homogenous coordinates also.

## 4 Differences between camera resection and the PnP problem

The main difference is that in the PnP problem, the camera is a SAME calibrated one, but in the camera resection, the camera is unspecified and allowed to change. Such a difference implies that in the PnP problem, any pairs of control points must subtend a fixed angle with different perspective centers, but in the camera resection, it only requires that the pencils of projection rays with different perspective centers be related by a homography, or projectively equivalent, which is much less stringent.

**Twisted cubic.** Here we would like to have some words about the twisted cubic. As shown in [12, p. 57], a twisted cubic is defined to be a curve in 3D projective space given in parameter form as:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = A \begin{pmatrix} 1 \\ \theta \\ \theta^2 \\ \theta^3 \end{pmatrix} = \begin{pmatrix} a_{11} + a_{12}\theta + a_{13}\theta^2 + a_{14}\theta^3 \\ a_{21} + a_{22}\theta + a_{23}\theta^2 + a_{24}\theta^3 \\ a_{31} + a_{32}\theta + a_{33}\theta^2 + a_{34}\theta^3 \\ a_{41} + a_{42}\theta + a_{43}\theta^2 + a_{44}\theta^3 \end{pmatrix}$$

where $A$ is a non-singular $4 \times 4$ matrix, $\theta$ is the parameter.

A twisted cubic can be uniquely determined by at least 7 control points, of which no four of them are coplanar. It is proved that when the control points lie on a twisted cubic, no matter how many control points we have, the camera resection is always indeterminate. However, it is completely wrong to think the PnP problem must also be degenerate when the control points are on a twisted cubic. In fact, the opposite is true and a unique solution can be generally obtained.

In addition, when the number of point correspondences is less than 6, the camera resection is always indeterminate even if the control points lie in general position, however the corresponding PnP problem will in general have a limited number of solutions, and in most cases, a unique solution.

Finally, let us take an example to geometrically illustrate the differences, where all the control points and the perspective center are coplanar for simplicity. In this coplanar case, the indeterminacy occurs in the camera resection when all the control points and the perspective center lie on a conic, as shown in Fig. 2, but only when the control points and the perspective center lie on a circle, the PnP problem becomes indeterminate, as shown in Fig. 3.

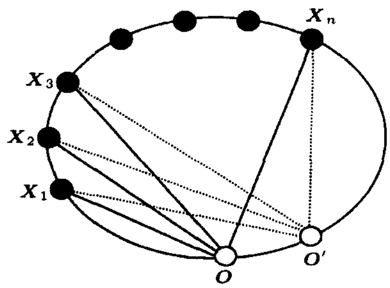


Fig. 2 The control points $\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_3, \cdots, \boldsymbol{X}_n$ project in equivalent ways to the two camera centers $\boldsymbol{O}$ and $\boldsymbol{O}'$. In other words, the perspective center $\boldsymbol{O}$ can be anywhere on the conic, the corresponding projection matrix is always a solution of the camera resection, and there are an infinite number of such solutions

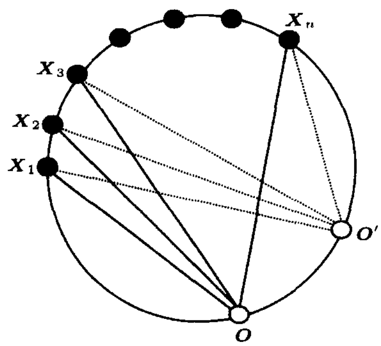


Fig. 3 Only when the control points and the perspective center lie on a circle, the PnP problem becomes indeterminate. In other words, the perspective center $\boldsymbol{O}$ can be anywhere on one of arc segments $\boldsymbol{X}_i\boldsymbol{X}_{1+\mathrm{mod}(i,n)}$ $i = 1, 2, 3 \cdots, n$ depending on the known angles $\angle \boldsymbol{X}_i\boldsymbol{O}\boldsymbol{X}_{1+\mathrm{mod}(i,n)}$ $i = 1, 2, 3, \cdots, n$ the corresponding distance set $(|\boldsymbol{OX}_1|, |\boldsymbol{OX}_2|, \cdots, |\boldsymbol{OX}_n|)$ will always be a solution to the PnP problem, and there are an infinite number of such solutions (Note: Arc segment $\boldsymbol{X}_i\boldsymbol{X}_{1+\mathrm{mod}(i,n)}$ denotes the segment from point $\boldsymbol{X}_i$ to $\boldsymbol{X}_{1+\mathrm{mod}(i,n)}$ as depicted in the Figure)

## References

1 Fishler M A, Bolles R C. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography, *Communications of the ACM*, 1981, **24**(6): 381~395

2 Haralick R M, Lee Chung-Nan, Karsten Ottenberg, Mivhael Noelle. Analysis and Solution of the Three Point Perspective Pose Estimation Problem, In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Maui, Hawaii: Medioni and Negahdaripour Eds., 1991, 592~598

3 Radu Horaud, Bernard Conio, Olivier Leboulleux. An Analytic Solution for the Perspective 4-Point Problem, Computer Vision, *Graphics and Image Processing*, 1989, **47**(1): 33~44

4 William J Wolfe, Donald Mathis. The perspective view of three points, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1991, **13**(1): 66~73

5 Abidi M A, Chandra T. A new efficient and direct solution for pose estimation using quadrangular targets: Algorithm and evaluation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1995, **17**(5): 534~538

6 Su Cheng, Xu Ying-Qing, Li Hua, Liu Shen-Quan. Necessary and Sufficient Condition of Positive Root Number of Perspective-Tree-Point Problem, *Chinese Journal of Computers*, 1998, **21**(12): 1084~1095

7 Quan L, Lan Z D. Linear N-point camera pose determination, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1999, **21**(8): 774~780

8 Fiore Paul D. Efficient linear solution of exterior orientation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, **23**(2): 140~148

9 Hu Z Y, Wu F C. A note on the number of solutions of the non-coplanar P4P problem, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, **24**(4): 550~555

10 Adnan Ansar, Kostas Daniilidis. Linear pose estimation from points and lines, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, **25**(5): 578~589

11 Gao X S, Hou X R, Tang J L, Cheng H F. Complete solution classification for the perspective-three-point problem, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, **25**(8): 930~943

12 Hartely R, Zisserman A. Multiple View Geometry in Computer Vision, Cambridge University Press, London, 2000

13 Wu Yi-Hong, Hu Zhan-Yi. Perspective-N-Point Camera Pose Determination under Different Definitions, Technical Report, NLPR-RV-03-014, National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, 2003

**SUN Feng-Mei** Received her master degree in the Department of Physics at the Northest Normal University in 1982. Since then, she has been with the North China University of Technology where she is now an associate professor. Her research interests include photo-electric information processing and 3D computer vision.