Journal of Mathematical Imaging and Vision 24: 131–141, 2006

DOI: 10.1007/s10851-005-3617-z

PnP Problem Revisited

YIHONG WU AND ZHANYI HU
National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, P.O. Box 2728, Beijing 100080, People's Republic of China
yhwu@nlpr.ia.ac.cn
huzy@nlpr.ia.ac.cn

Published online: 7 December 2005

Abstract. Perspective-n-Point camera pose determination, or the PnP problem, has attracted much attention in the literature. This paper gives a systematic investigation on the PnP problem from both geometric and algebraic standpoints, and has the following contributions: Firstly, we rigorously prove that the PnP problem under distance-based definition is equivalent to the PnP problem under orthogonal-transformation-based definition when $n > 3$, and equivalent to the PnP problem under rotation-transformation-based definition when $n = 3$. Secondly, we obtain the upper bounds of the number of solutions for the PnP problem under different definitions. In particular, we show that for any three non-collinear control points, we can always find out a location of optical center such that the P3P problem formed by these three control points and the optical center can have 4 solutions, its upper bound. Additionally a geometric way is provided to construct these 4 solutions. Thirdly, we introduce a depth-ratio based approach to represent the solutions of the whole PnP problem. This approach is shown to be advantageous over the traditional elimination techniques. Lastly, degenerated cases for coplanar or collinear control points are also discussed. Surprisingly enough, it is shown that if all the control points are collinear, the PnP problem under distance-based definition has a unique solution, but the PnP problem under transformation-based definition is only determined up to one free parameter.

Keywords: perspective-n-point camera pose determination, distance-based definition, transformation-based definition, depth-ratio based equation, upper bound of the number of solutions

1. Introduction

The aim of Perspective-n-Point camera pose determination, popularly called the PnP problem, is to determine the relative position between camera and scene from n known correspondences of space control points and image points. The PnP problem finds its important applications in computer vision and photogrammetry, and has attracted much attention since its formal introduction in 1981, for example [1–15] to cite a few.

In the past, the PnP problem was known under two different definitions, the distance-based definition and transformation-based definition [11]. Different definition usually implies different set of solutions, different computational complexity, and different robustness. The distance-based definition was first formally introduced by Fishler and Bolles in 1981 [4], and the transformation-based definition was given by Horaud, Conio, and Leboulleux in 1989 [10]. Then, the PnP problem has been extensively studied in the literature under these two definitions. In 2002, Hu and Wu [11] noted that these two definitions are in general not equivalent by a specific example of non-coplanar P4P problem. What on earth is the difference between these two definitions for the whole PnP problem? In Section 3 of this paper, by rigorous mathematical derivations, we present the relationships of the PnP problem under different definitions and the converting ways of solutions between different definitions. Since under different definitions, the associated computational complexities are usually different, we can solve the PnP problem by selecting an appropriate definition, then compute solutions under other definitions by the relationships and converting ways presented in this work (Section 3).

In Hu and Wu [11], an algebraic method for proving the upper bound of the number of solutions for the P4P problem is presented. What do the configurations for the multiple solutions look like? Can we have a geometric proof for the upper bounds of the number of solutions for the whole PnP problem? In Section 4, from a geometric standpoint, we obtain the upper bounds of the PnP problem under different definitions, and provide some insights into why different definitions usually have different bounds. In particular, for the P3P problem, we show that given 3 arbitrary non-collinear control points, we can always find out an optical center such that the P3P problem formed by these 3 control point and the optical center will have 4 solutions, i.e., the P3P problem can attain its upper bound of the number of solutions (Section 4).

Quan and Lan [12] gave a linear SVD method to solve the P4P and P5P problems. Before applying SVD, univariate equations must be obtained from the nonlinear equations by resultant eliminations. Here we propose a depth-ratio based technique to simply and directly represent the elimination results. By a combination of our depth-ratio based elimination and their SVD based method, the PnP problem can be solved more efficiently. Gao et al. [6] gave a complete solution set for the P3P problem by a algebraic way. Our work is rather geometric and more instructive. Of course, at this stage, in contrast to their work, we have not achieved a complete solution set (Section 5).

Some degenerated cases, i.e., the coplanar control points, or collinear control points, are also investigated. It is shown that if all the control points are collinear, the PnP problem under distance-based definition has a unique solution, but the solution set of the PnP problem under transformation-based definition is infinitely large (Section 6).

Prior to any further discussions, we would at first point out:

(1) Due to the limited space, except for Result 3, we do not give proofs for other results, the interested reader could find them in [15].
(2) Since the PnP problem is under constrained when n <3, and can be linearly determined when $n \geq 6$, we always assume $n = 3$, 4, 5 in this work.

2. Preliminaries

In Sections 3, 4, 5, we always assume that no three of control points are collinear, no four of control points are coplanar, and the optical center is not coplanar with any three-control points.

2.1. *Notation*

A bold capital letter denotes a matrix or a non-homogeneous 3-vector, a capped bold capital letter such as $\hat{\mathbf{M}}$, denotes the homogeneous 4-vector for the non-homogeneous 3-vector $\mathbf{M}$, a bold small letter $\mathbf{m}$ denotes homogeneous image points, and "≈" denotes the equality up to a scalar.

A 3×3 matrix $\mathbf{R}$ is an orthogonal matrix if $\mathbf{R}\,\mathbf{R}^{\tau} = I$ and $\mathbf{R}^{\tau}\mathbf{R} = \mathbf{I}$ where $\mathbf{I}$ is the 3×3 identical matrix. Moreover, $\mathbf{R}$ is a rotation matrix if $\det(\mathbf{R}) = 1$ and a reflection matrix if $\det(\mathbf{R}) = -1$.

2.2. *Two Definitions and Their Constraint Equations*

The distance-based definition of the PnP problem is defined as [4]:

> Given the relative spatial locations of n points $\mathbf{M}_i$, $i = 1, \ldots, n$, and given the angle $\angle\mathbf{M}_i\mathbf{O}\mathbf{M}_j$ to every pair of these points from the perspective center $\mathbf{O}$ (the camera's optical center), find the lengths $Xi = |\mathbf{O}\mathbf{M}_i|$ of the line segments joining the perspective center to each of these points.

The constraint equations under this definition, called Definition 1, are:

$$X_i^2 + X_j^2 - 2X_iX_j\cos\theta_{ij} = d_{ij}^2, \quad i \neq j \quad (2.1)$$

where x_i, x_j are variables, $\theta_{ij} = \angle\mathbf{M}_i\mathbf{O}\mathbf{M}_j$ and $d_{ij} = |\mathbf{M}_i\mathbf{M}_j|$ (the distance between $\mathbf{M}_i$ and $\mathbf{M}_j$) are known parameters.

The transformation-based definition of the PnP problem is defined as [10]:

> Given n points $\mathbf{M}_i, = 1, \ldots, n$ with known coordinates in an object centered frame and their corresponding projections $\mathbf{m}_i$, $i = 1, \ldots, n$ onto an image plane, and given the intrinsic camera parameters $\mathbf{K}$, find the transformation (a rotation matrix $\mathbf{R}$ and a translation vector $\mathbf{t}$) between the object frame and the camera frame.

The constraint equations under this definition, called Definition 2, are:

$$\mathbf{m}_i \approx \mathbf{K}(\mathbf{R}, \mathbf{t})\hat{\mathbf{M}}_i, \quad \text{with } \mathbf{R} \text{ a rotation matrix} \qquad (2.2)$$

where $\mathbf{R}$ and $\mathbf{t}$ are variables, $\mathbf{K}$, $\mathbf{M}_i$, and $\mathbf{m}_i$ are known parameters.

In Definitions 1 and 2, the points $\mathbf{M}_i$, $i = 1 \dots n$ are called control points.

3. Relationships of the PnP Problem Under Different Definitions

Definition 1 and Definition 2 are generally not equivalent as shown in [11] by a specific example. What on earth are the differences between these two definitions? Under what conditions are the two definitions equivalent? How to convert the solutions from one definition to another? In this section, these questions will be answered through rigorous mathematical derivations.

For the convenience of subsequent discussions, we set up a new definition, called the orthogonal-transformation-based definition for the PnP problem as:

> Given n points $\mathbf{M}_i, i = 1, \dots, n$ with known coordinates in an object centered frame and their corresponding projections $\mathbf{m}_i, i = 1, \dots, n$ onto an image plane, and given the intrinsic camera parameters $\mathbf{K}$, find the orthogonal transformation (an orthogonal matrix $\mathbf{R}$ and a translation vector $\mathbf{t}$) between the object frame and the camera frame.

The constraint equations under this definition, called Definition 3 are:

$$\mathbf{m}_i \approx \mathbf{K}(\mathbf{R}, \mathbf{t})\hat{\mathbf{M}}_i, \quad \text{with } \mathbf{R} \text{ a rotation matrix} \qquad (3.1)$$

where $\mathbf{R}$ and $\mathbf{t}$ are variables, and $\mathbf{K}$ and $\mathbf{M}_i$, $\mathbf{m}_i$ are known parameters.

Evidently when $\mathbf{R}$ in (3.1) is a rotation matrix, then Definition 3 becomes Definition 2, therefore, Definition 3 includes Definition 2. In order to avoid possible confusions, the original transformation-based definition Definition 2 is called the rotation-transformation-based definition in the following.

(3.1) can be changed to:

$$x_i \mathbf{m}_i = \mathbf{K}(\mathbf{R}, \mathbf{t})\hat{\mathbf{M}}_i$$

where x_i is an unknown scalar, $\mathbf{R}$ an orthogonal matrix. Because for the PnP problem, none of the n control points is at infinity, this equation can be further changed into:

$$x_i \mathbf{m}_i = \mathbf{K}(\mathbf{R}\mathbf{M}_i + \mathbf{t}), \quad \text{or} \quad x_i \mathbf{K}^{-1}\mathbf{m}_i = \mathbf{R}\mathbf{M}_i + \mathbf{t}$$

Since $\mathbf{K}$ and $\mathbf{m}_i$ are known, $\mathbf{K}^{-1}\mathbf{m}_i$, called the normalized image, is also known. Moreover, because no image point is at infinity, we can always homogenize $\mathbf{K}^{-1}\mathbf{m}_i$ such that its last coordinate is 1. Thus, in sequel, $\mathbf{m}_i$ is always referred to the normalized image points with 1 as its last coordinate. Then the above equation becomes:

$$x_i \mathbf{m}_i = \mathbf{R}\mathbf{M}_i + \mathbf{t} \qquad (3.2)$$

It is shown that $\mathbf{M}_i$ lies in front of the camera if and only if the depth $\det(\mathbf{R})x_i > 0$ [9, p. 501], hence if $\mathbf{R}$ is a rotation matrix, $x_i > 0$, and $x_i < 0$ if $\mathbf{R}$ is a reflection matrix.

The distance from the optical center to each control point becomes infinitely large, and the PnP problem under Definition 1 becomes meaningless if the optical center $\mathbf{O}$ is at infinity. So we assume that the camera optical center is not at infinity, at the time $\mathbf{O} = -\mathbf{R}^T\mathbf{t}$, then (3.2) can be rewritten as:

$$x_i \mathbf{R}^\tau \mathbf{m}_i = \mathbf{M}_i - \mathbf{O} \qquad (3.3)$$

By (3.3), the following Result 1 can be obtained.

Result 1. With notations as before, from (3.2) we have:

$$\begin{cases} X_i = x_i\sqrt{\mathbf{m}_i \cdot \mathbf{m}_i}, & \text{when } x_i > 0 \\ X_i = -x_i\sqrt{\mathbf{m}_i \cdot \mathbf{m}_i}, & \text{when } x_i < 0 \end{cases}.$$

and

$$\cos < \mathbf{M}_i \mathbf{O} \mathbf{M}_j = \frac{\mathbf{m}_i \cdot \mathbf{m}_j}{\sqrt{\mathbf{m}_i \cdot \mathbf{m}_i}\sqrt{\mathbf{m}_j \cdot \mathbf{m}_j}}$$

By (3.2), for $i \neq j = 1..n$, we have:

$$x_i \mathbf{m}_i - x_j \mathbf{m}_j = \mathbf{R}(\mathbf{M}_i - \mathbf{M}_j)$$

Since $\mathbf{R}$ is an orthogonal matrix, $\mathbf{R}^\tau\mathbf{R} = \mathbf{I}$, there is:

$$(x_i \mathbf{m}_i - x_j \mathbf{m}_j) \cdot (x_i \mathbf{m}_i - x_j \mathbf{m}_j) = d_{ij}^2 \qquad (3.4)$$

Because all control points lie in front of the camera, if one of x_i, $i = 1, \dots, n$, is positive (negative), then all of them are positive (negative) for a group of solutions of x_i, $i = 1,\ldots,n$ in (3.4). Thus by Result 1, (3.4) is changed to:

$$\begin{aligned} d_{ij}^2 &= (x_i\mathbf{m}_i - x_j\mathbf{m}_j)\cdot(x_i\mathbf{m}_i - x_j\mathbf{m}_j) \\ &= x_i^2\mathbf{m}_i\cdot\mathbf{m}_i - 2x_ix_j\mathbf{m}_i\cdot\mathbf{m}_j + x_j^2\mathbf{m}_j\cdot\mathbf{m}_j \\ &= X_i^2 + X_j^2 - 2X_iX_j\cos\theta_{ij} \end{aligned}$$

It is just the Eq. (2.1) under Definition 1. It follows that the determination of the positive solutions of X_i by (2.1) is equivalent to determining the solutions of x_i by (3.4) with all control points lying in front of the camera.

It is clear that if x_i, $i = 1\ldots n$, is a group of solutions of (3.4), so is $-x_i$, $i = 1\ldots n$. Thus we need only consider the positive solutions of (3.4), which correspond to the case that all control points lie in front of the camera when $\mathbf{R}$ is a rotation matrix, and all the control points are behind the camera when $\mathbf{R}$ is a reflection matrix. Based on the above reasoning, we can conclude that to determine the positive solutions of X_i by (2.1) is equivalent to determining the positive solutions of x_i by (3.4).

Result 2. From (3.4), for three different i, j, k, we have:

$$\begin{aligned} &(x_i\mathbf{m}_i - x_k\mathbf{m}_k)\cdot(x_j\mathbf{m}_j - x_k\mathbf{m}_k) \\ &\quad= \frac{d_{ik}^2 + d_{jk}^2 - d_{ij}^2}{2} \\ &\quad= d_{ik}d_{jk}\cos < \mathbf{M}_i\mathbf{M}_k\mathbf{M}_j \\ &\quad= (\mathbf{M}_i - \mathbf{M}_k)\cdot(\mathbf{M}_j - \mathbf{M}_k) \end{aligned}$$

Without loss of generality, we set up the object-centered frame as shown in Fig. 1, where $\mathbf{M}_1 = (0,0,0)^\tau$, $\mathbf{M}_2 = (d_{12},0,0)^\tau$, $\mathbf{M}_3 = (a,b,0)^\tau$ with $b > 0$. Then (3.2) becomes:

$$\begin{cases} x_1\mathbf{m}_1 = \mathbf{t} \\ x_2\mathbf{m}_2 = d_{12}\mathbf{r}_1 + \mathbf{t} \\ x_3\mathbf{m}_3 = a\mathbf{r}_1 + b\mathbf{r}_2 + \mathbf{t} \end{cases} \tag{3.5}$$

Under this frame and Result 2, we have the following theorem.

Theorem 1. *The P3P problem under distance-based definition (Definition 1) is equivalent to the P3P problem under rotation-transformation-based definition (Definition 2).*

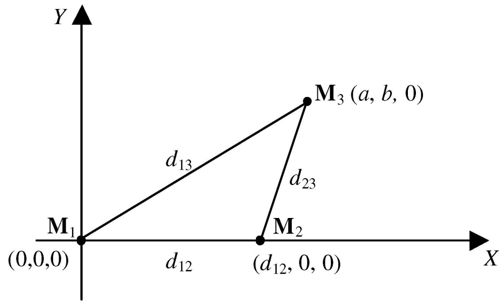

Figure 1. The *XY*-plane for the object frame.

Remark 1. A group of solutions of (x_1, x_2, x_3) by (3.4) corresponds to a group of solutions of $(\mathbf{r}_1, \mathbf{r}_2, \mathbf{t})$ by (3.5). From $\mathbf{r}_1$ and $\mathbf{r}_2$, a rotation matrix $\mathbf{R} = [\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_1\times\mathbf{r}_2]$ and a reflection matrix $\mathbf{R}' = [\mathbf{r}_1, \mathbf{r}_2, -\mathbf{r}_1\times\mathbf{r}_2]$ can be formed. However, only $(\mathbf{R}, \mathbf{t})$ is a group of solutions under Definition 2, but both $(\mathbf{R}, \mathbf{t})$ and $(\mathbf{R}', \mathbf{t})$ are groups of solutions under Definition 3. Thus, one group of solutions of (x_1, x_2, x_3) corresponds to one group of solutions under Definition 2, and two groups of solutions under Definition 3. The two optical centers associated with $(\mathbf{R}, \mathbf{t})$ and $(\mathbf{R}', \mathbf{t})$ are:

$$\mathbf{O} = -\mathbf{R}^\tau\mathbf{t} = \begin{pmatrix} -\mathbf{r}_1^\tau\mathbf{t} \\ -\mathbf{r}_2^\tau\mathbf{t} \\ -(\mathbf{r}_1\times\mathbf{r}_2)^\tau\mathbf{t} \end{pmatrix},$$

$$\mathbf{O}' = -\mathbf{R}'^\tau\mathbf{t} = \begin{pmatrix} -\mathbf{r}_1^\tau\mathbf{t} \\ -\mathbf{r}_2^\tau\mathbf{t} \\ (\mathbf{r}_1\times\mathbf{r}_2)^\tau\mathbf{t} \end{pmatrix}$$

That is, the two optical centers $\mathbf{O}$ and $\mathbf{O}'$ are symmetric with respect to *XY*-plane, or the plane $\mathbf{M}_1\mathbf{M}_2\mathbf{M}_3$.

As shown in [15], we have the following theorem for the PnP problem with $n \geq 4$.

Theorem 2. *When $n \geq 4$, the PnP problem under distance-based definition Definition 1 is equivalent to the PnP problem under orthogonal-transformation-based definition Definition 3.*

Theorems 1 and 2 clarify the relationships of the PnP problem under different definitions. Such relationships are summarized in Fig. 2 for the P3P problem and in Fig. 3 for the PnP problem with $n > 3$.

Solutions under different definitions can be converted to each other as:

$$X_i \xleftrightarrow{X_i = x_i\sqrt{\mathbf{m}_i\cdot\mathbf{m}_i}} x_i \xleftrightarrow{x_i\,\mathbf{m}_i = \mathbf{R}\,\mathbf{M}_i + \mathbf{t}} \mathbf{R}, \mathbf{t}$$

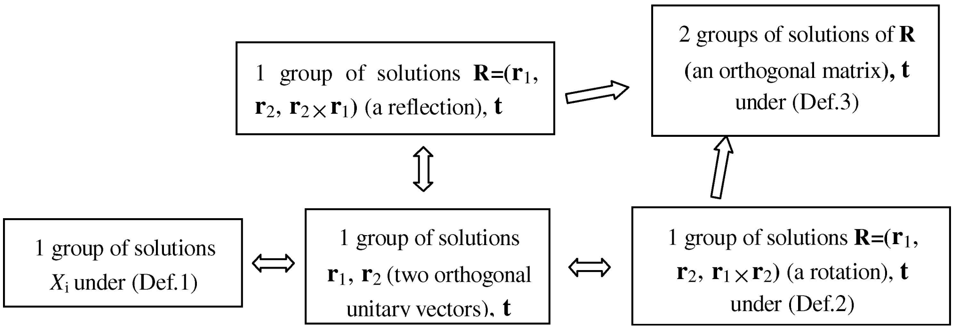

Figure 2. The relationships of the P3P problem under different definitions.

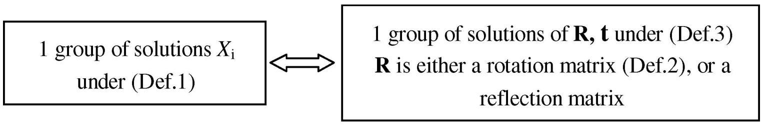

Figure 3. The relationships of the PnP problem under different definitions with $n > 3$.

The details can be found in [15].

4. Upper Bounds of the Number of Solutions

In this section, we give upper bounds under different definitions from a geometric way.

By Bezout theorem, we know that the equations in (3.4) for the P3P problem have at most 8 solutions. When x_i, $i = 1, 2, 3$, is a group of solutions, so is $-x_i$, $i = 1, 2, 3$. Hence there are at most 4 groups of positive solutions of x_i. Moreover, since each group of positive solutions x_i corresponds to one group of positive solutions X_i and to one group of solutions ($\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{t}$), and vice versa, the P3P problem has at most 4 groups of solutions under both Definitions 1 and 2, and 8 groups of solutions under Definition 3 (see Remark 1 in Section 3). The following Theorem 3, or [4, 11] show that these upper bounds can also be attainable.

Besides, in the literature [4, 11], the attainable upper bound of the P3P problem is always illustrated by the example: the three control points forms equilateral triangle. In fact, the 3 control points could be any three non-collinear ones.

Theorem 3. *For any 3 non-collinear control points, there always exist an optical center such that the P3P problem formed by these 3 control points and the optical center has 4 solutions under (Definition 1) or (Definition 2), and 8 solutions under Definition 3.*

The interested reader is referred to [15] for the proof. The following is a sketch of it:

Note that since the PnP problem is to determine the relative pose between the camera and objects, for convenience purpose, we can here let the camera fixed and the objects varied, such setting is rather unusual compared with the traditional setting where the camera is assumed varied and the objects fixed.

Case 1. The triangle $\mathbf{M}_1\mathbf{M}_2\mathbf{M}_3$ is not a right-angled one.

Let $\mathbf{G}$ be the orthocenter of the triangle and the optical center $\mathbf{O}$ be far from the triangle. If $\mathbf{OG}$ is perpendicular to the plane $\mathbf{M}_1\mathbf{M}_2\mathbf{M}_3$, then the four solutions can be constructed as follows:

Rotate the triangle with the axis $\mathbf{M}_2\mathbf{M}_3$. As shown in Fig. 4, since $\mathbf{M}_2\mathbf{M}_3$ is orthogonal to $\mathbf{OM}_1$, the locus of $\mathbf{M}_1$ must intersect the line $\mathbf{OM}_1$ at another point, denoted by $\mathbf{M}'_1$. Similarly, rotate the triangle with the axis $\mathbf{M}_1\mathbf{M}_3$ and $\mathbf{M}_1\mathbf{M}_2$, we can construct a point $\mathbf{M}'_2$ on $\mathbf{OM}_2$ and a point $\mathbf{M}'_3$ on $\mathbf{OM}_3$. Then the four positive solutions of the P3P problem are: $\{\mathbf{O}, (\mathbf{M}_1\mathbf{M}_2\mathbf{M}_3)\}$, $\{\mathbf{O}, (\mathbf{M}'_1\mathbf{M}_2\mathbf{M}_3)\}$, $\{\mathbf{O}, (\mathbf{M}_1\mathbf{M}'_2\mathbf{M}_3)\}$ and $\{\mathbf{O}, (\mathbf{M}_1\mathbf{M}_2\mathbf{M}'_3)\}$. The condition that $\mathbf{O}$ is far away from the plane $\mathbf{M}_1\mathbf{M}_2\mathbf{M}_3$ ensures that $\mathbf{M}_i$ and $\mathbf{M}'_i$ lie in the same side of $\mathbf{O}$ on the line $\mathbf{OM}_i$.

Case 2. The triangle is a right-angled one, say $\angle\mathbf{M}_2 = 90°$.

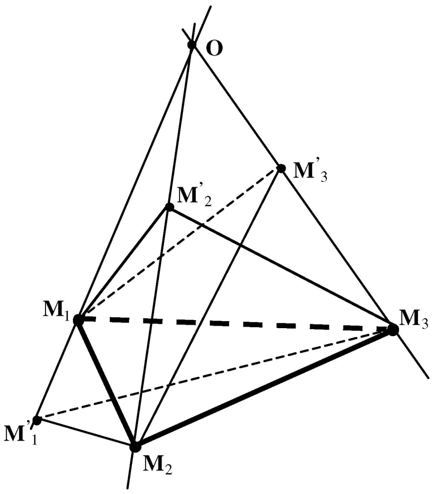

Figure 4. The 4 solutions for a non-right-angled triangle.

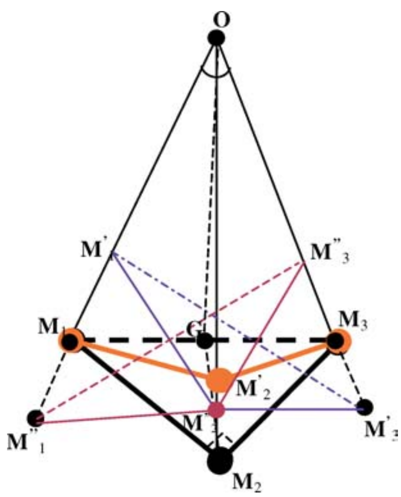

Figure 5. The 4 solutions for a right-angled triangle.

Let $\mathbf{G}$ be the normal foot at $\mathbf{M}_1\mathbf{M}_3$ of $\mathbf{M}_2$, if $\mathbf{OG}$ is perpendicular to the plane $\mathbf{M}_1\mathbf{M}_2\mathbf{M}_3$, then the four solutions shown in Fig. 5 are: $\{\mathbf{O}, (\mathbf{M}_1\mathbf{M}_2\mathbf{M}_3)\}$, $\{\mathbf{O}, (\mathbf{M}_1\mathbf{M}'_2\mathbf{M}_3)\}$, $\{\mathbf{O}, (\mathbf{M}'_1\mathbf{M}''_2\mathbf{M}'_3)\}$ and $\{\mathbf{O}, (\mathbf{M}''_1\mathbf{M}''_2\mathbf{M}''_3)\}$, where $\mathbf{M}'_2$ is constructed similarly as in Case 1, but $\mathbf{M}''_1$, $\mathbf{M}''_2$, $\mathbf{M}''_3$ are found by some algebraic method. At this moment, we cannot yet give a direct geometric construction for the latter two solutions.

Now let us proceed for the PnP problem with $n > 3$. Take the coordinate system as that in Fig. 1. Suppose that the upper bounds are attainable for the P3P problem, let $(\mathbf{R}_i = (\mathbf{r}_{1i}, \mathbf{r}_{2i}, \mathbf{r}_{1i} \times \mathbf{r}_{2i}), \mathbf{t}_i)$, and $(\mathbf{R}'_i = (\mathbf{r}_{1i}, \mathbf{r}_{2i}, -\mathbf{r}_{1i} \times \mathbf{r}_{2i}), \mathbf{t}_i)$, $i = 1\ldots4$ be the 8 groups of solutions under Definition 3. Now add other control points $\mathbf{M}_i$, $i = 4\ldots n$, $n \geq 4$ to the P3P problem, then the possible solutions of the PnP problem under Definitions 2 and 3 must be among $(\mathbf{R}_i, \mathbf{t}_i)$ and $(\mathbf{R}'_i, \mathbf{t}_i)$. It follows that the number of solutions for the PnP problem under Definitions 2 or 3 will be no more than that for the P3P problem, i.e. 4 or 8. Moreover by Theorem 2, its number of solutions under Definition 1 will also be no more than 8.

Result 3. Out of the four pairs of solutions $((\mathbf{R}_i, \mathbf{t}_i), (\mathbf{R}'_i, \mathbf{t}_i))$, $i = 1\ldots4$, of the P3P problem, there exists at most one pair $((\mathbf{R}_{i_0}, t_{i_0}), (\mathbf{R}'_{i_0}, t_{i_0}))$ such that both $(\mathbf{R}_{i_0}, t_{i_0})$ and $(\mathbf{R}'_{i_0}, t_{i_0})$ are the solutions of the PnP problem, $n \geq 4$. If there exists indeed such a pair $((\mathbf{R}_{i_0}, t_{i_0}), (\mathbf{R}'_{i_0}, t_{i_0}))$, then:

(i) The optical centers $\mathbf{O}_{i_0} = -\mathbf{R}^{\tau}_{i_0}\mathbf{t}_{i_0}$, $\mathbf{O}'_{i_0} = -\mathbf{R}'^{\tau}_{i_0}\mathbf{t}_{i_0}$, and $\mathbf{M}_i$, $i = 4\ldots n$ should be collinear, and the image points of $\mathbf{M}_i$, $i = 4\ldots n$ are all coincident.
(ii) The PnP problem has only two solutions $(\mathbf{R}_{i_0}, \mathbf{t}_{i_0})$ and $(\mathbf{R}'_{i_0}, \mathbf{t}_{i_0})$ with $n \geq 5$.

The proof of Result 3 is given in Appendix.
Therefore, we have the following theorem:

Theorem 4.

(1) The number of the solutions of the P4P problem is at most 4 under Definition 2, at most 5 under both Definition 1 and 3.
(2) The number of the solutions of the P5P problem is at most 4 under all three definitions.

Figure 6 is an illustration for the upper bounds: $\mathbf{O}_i = -\mathbf{R}^{\tau}_i\mathbf{t}_i$, $\mathbf{O}'_i = -\mathbf{R}'^{\tau}_i\mathbf{t}_i$, $i = 1\ldots4$ are the 4 pairs of symmetric optical centers of the P3P problem to the plane $\mathbf{M}_1\mathbf{M}_2\mathbf{M}_3$. With other control points $\mathbf{M}_i$, $i = 4\ldots n$ added, the optical centers of the solutions of the PnP problem $(n > 3)$ should be among those $\mathbf{O}_i$, $\mathbf{O}'_i$ $i = 1\ldots4$. Because $\mathbf{O}_i$ $i = 1\ldots4$ are the optical centers of the solutions under Definition 2 and $(\mathbf{O}_i, \mathbf{O}'_i)$ $i = 1\ldots4$ are those under Definition 3, and *the PnP problem* $(n = 4, 5)$ *has at most 4 solutions under Definition 2*. Moreover, by Result 3, there could at most exist one pair of $(\mathbf{O}_i, \mathbf{O}'_i)$ as the solutions of the optical centers, so *the P4P problem has at most 5 solutions under Definition 3*. If the upper bound of 5 is attained, there must exist such a pair as the solutions, and $\mathbf{M}_4$ must be on the line through the pair of optical centers. If

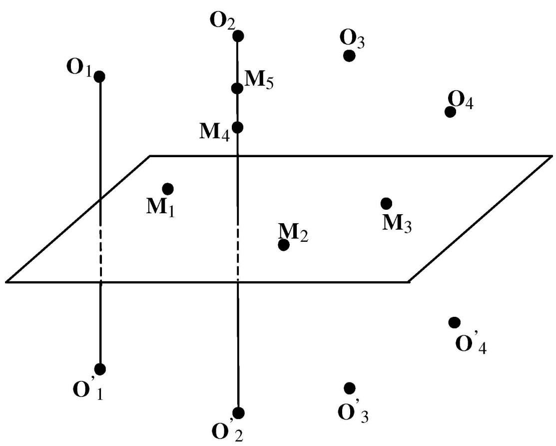

Figure 6. Optical centers and control points for multiple solutions, if there is a pair of the symmetric optical centers as the solutions, say as **O**2, $\mathbf{O}'_2$, then $\mathbf{M}_4$, $\mathbf{M}_5$ are on the line through them. At the time, for the P4P problem, there are at most 5 solutions; for the P5P problem, there are only these two solutions. Where if does not exist such a pair as the solutions, then both P4P and P5P problem have at most 4 solutions.

such a pair does exist, by Result 3, the P5P problem has only two solutions under Definition 3, and their optical centers must be this pair. Else, the P5P problem has at most 4 solutions under Definition 3. In summary, *the upper bound of the P5P problem is 4 under Definition 3*. Since Definitions 1 and 3 are equivalent (Theorem 2) for the P4P problem and P5P problem, their upper bounds under Definition 1 are the same as those under Definition 3.

5. Quasi-Analytical Solutions of the PnP Problem

Equations (2.1) under Definition 1 are cyclic, so there would be some difficulties if we directly solve it by symbolic method. It follows that many different techniques were introduced in the literature. Here, we will give a simpler method to directly represent the PnP problem by an equivalent system of univariate equations. Since only a system of univariate equations, rather than explicit solutions, is provided in this work, the solutions are conventionally dubbed "quasi-analytical".

Rather than (2.1), (3.4) is used here.

Since $x_1 \neq 0$, by dividing both (3.4) and the equations in Result 2 from both sides by x_1^2, we have:

$$\begin{cases} (y_i\mathbf{m}_i - y_j\mathbf{m}_j)\cdot(y_i\mathbf{m}_i - y_j\mathbf{m}_j) = \dfrac{d_{ij}^2}{x_1^2}, \\ (y_i\mathbf{m}_i - y_k\mathbf{m}_k)\cdot(y_j\mathbf{m}_j - y_k\mathbf{m}_k) = \dfrac{d_{ik}^2 + d_{jk}^2 - d_{ij}^2}{2x_1^2}, \end{cases}$$
$$i, j, k = 1..n, \quad i \neq j \neq ks \tag{5.1}$$

where $y_i = \frac{x_i}{x_1}$, $i = 1\ldots n$ (note that $y_1 = 1$). Since $y_2\mathbf{m}_2 - \mathbf{m}_1 \neq 0$ (otherwise $\mathbf{m}_1 \approx \mathbf{m}_2$), $(y_2\mathbf{m}_2 - \mathbf{m}_1)\cdot(y_2\mathbf{m}_2 - \mathbf{m}_1) \neq 0$ follows. So by $(y_2\mathbf{m}_2 - \mathbf{m}_1)\cdot(y_2\mathbf{m}_2 - \mathbf{m}_1) = \frac{d_{12}^2}{x_1^2}$, there is:

$$x_1^2 = \frac{d_{12}^2}{(y_2\mathbf{m}_2 - \mathbf{m}_1)\cdot(y_2\mathbf{m}_2 - \mathbf{m}_1)} \tag{5.2}$$

Substituting it into the equations in (5.1), we have:

$$\begin{cases} (y_i\mathbf{m}_i - y_j\mathbf{m}_j)\cdot(y_i\mathbf{m}_i - y_j\mathbf{m}_j) = \dfrac{d_{ij}^2}{d_{12}^2} \\ \quad\cdot(y_2\mathbf{m}_2 - \mathbf{m}_1)(y_2\mathbf{m}_2 - \mathbf{m}_1), \ (i, j) \neq (2, 1) \\ (y_i\mathbf{m}_i - y_k\mathbf{m}_k)\cdot(y_j\mathbf{m}_j - y_k\mathbf{m}_k) \\ \quad = \dfrac{d_{ik}^2 + d_{jk}^2 + d_{ij}^2}{2d_{12}^2}(y_2\mathbf{m}_2 - \mathbf{m}_1)\cdot(y_2\mathbf{m}_2 - \mathbf{m}_1) \end{cases} \tag{5.3}$$

It is clear that once a group of positive solutions of y_i, $i = 2\ldots n$ is determined by (5.3), the positive solution of x_1 can be uniquely determined by (5.2), and the positive solutions of x_i, $i = 2\ldots n$, can also be uniquely determined by $y_i = \frac{x_i}{x_1}$. Therefore, in the following, we are limited only to discuss the solutions of y_i, $i = 2\ldots n$ by (5.3).

Discard the dependent equations in (5.3) (see [15]), the equivalent form of (5.3) is:

$$\begin{cases} (y_i\mathbf{m}_i - \mathbf{m}_1)\cdot(y_i\mathbf{m}_i - \mathbf{m}_1) \\ \quad = \dfrac{d_{1i}^2}{d_{12}^2}(y_2\mathbf{m}_2 - \mathbf{m}_1)\cdot(y_2\mathbf{m}_2 - \mathbf{m}_1), \\ (y_i\mathbf{m}_i - \mathbf{m}_1)\cdot(y_j\mathbf{m}_j - \mathbf{m}_1) \\ \quad = \dfrac{d_{1i}^2 + d_{1j}^2 + d_{ij}^2}{d_{12}^2}(y_2\mathbf{m}_2 - \mathbf{m}_1)\cdot(y_2\mathbf{m}_2 - \mathbf{m}_1), \end{cases}$$
$$i = 3..n, \quad j = 2..i \neq j \tag{5.4}$$

5.1. *Quasi-Analytical Solutions of the P3P Problem*

When $n = 3$, there are two equations in (5.4) as:

$$\begin{cases} (y_3\mathbf{m}_3 - \mathbf{m}_1)\cdot(y_2\mathbf{m}_2 - \mathbf{m}_1) \\ \quad = \dfrac{d_{12}^2 + d_{13}^2 + d_{23}^2}{2d_{12}^2}(y_2\mathbf{m}_2 - \mathbf{m}_1)\cdot(y_2\mathbf{m}_2 - \mathbf{m}_1), \\ (y_3\mathbf{m}_3 - \mathbf{m}_1)\cdot(y_3\mathbf{m}_3 - \mathbf{m}_1) \\ \quad = \dfrac{d_{13}^2}{d_{12}^2}(y_2\mathbf{m}_2 - \mathbf{m}_1)\cdot(y_2\mathbf{m}_2 - \mathbf{m}_1) \end{cases} \tag{5.5}$$

The solutions can be classified as the following four cases. In each case, we directly represent the eliminated univariate equation system and give the corresponding classification conditions as well as their geometric meanings. Here a summary is merely provided, the more details can be found in [15].

Case (i): When $\mathbf{m}_3\cdot(y_2\mathbf{m}_2 - \mathbf{m}_1) \neq 0$, (5.5) can be expressed as:

$$\begin{cases} y_3 = \dfrac{\left(\left(d_{12}^2 + d_{13}^2 + d_{23}^2\right) y_2\mathbf{m}_2 + \left(d_{12}^2 - d_{13}^2 + d_{23}^2\right)\mathbf{m}_1\right)\cdot(y_2\mathbf{m}_2 - \mathbf{m}_1)}{2d_{12}^2\mathbf{m}_3\cdot y_2\mathbf{m}_2 - \mathbf{m}_1}, \\ (y_2\mathbf{m}_2 - \mathbf{m}_1)\cdot(y_2\mathbf{m}_2 - \mathbf{m}_1) = \dfrac{d_{12}^2}{d_{13}^2}(y_3\mathbf{m}_3 - \mathbf{m}_1)\cdot(y_3\mathbf{m}_3 - \mathbf{m}_1) \end{cases} \tag{C1}$$

By (3.3), the geometric meaning of $\mathbf{m}_3\cdot(y_2\mathbf{m}_2 - \mathbf{m}_1) \neq 0$ is that $\mathbf{OM}_3$ is not perpendicular to line $\mathbf{M}_1\mathbf{M}_2$.

Case (ii): When $\mathbf{m}_3\cdot(y_2\mathbf{m}_2 - \mathbf{m}_1) = 0$ but $\mathbf{m}_3\cdot\mathbf{m}_2 \neq 0$, the solution can be obtained by:

$$\begin{cases} y_2 = \dfrac{\mathbf{m}_1\cdot\mathbf{m}_3}{\mathbf{m}_2\cdot\mathbf{m}_3}, \\ (y_3\mathbf{m}_3 - \mathbf{m}_1)\cdot(y_3\mathbf{m}_3 - \mathbf{m}_1) \\ \quad = \dfrac{d_{13}^2}{d_{12}^2}(y_2\mathbf{m}_2 - \mathbf{m}_1)\cdot(y_2\mathbf{m}_2 - \mathbf{m}_1) \end{cases} \tag{C2}$$

The geometric meaning of the above conditions is:

$$\mathbf{OM}_3 \perp \mathbf{M}_1\mathbf{M}_2, \text{ and } \angle\mathbf{M}_1\mathbf{OM}_3 < 90^\circ, \text{ and } \angle\mathbf{M}_2\mathbf{OM}_3 < 90^\circ$$

or

$$\mathbf{OM}_3 \perp \mathbf{M}_1\mathbf{M}_2, \text{ and } \angle\mathbf{M}_1\mathbf{OM}_3 > 90^\circ, \text{ and } \angle\mathbf{M}_2\mathbf{OM}_3 > 90^\circ$$

Case(iii): When $\mathbf{m}_3\cdot(y_2\mathbf{m}_2 - \mathbf{m}_1) = 0$, and $\mathbf{m}_3\cdot\mathbf{m}_2 = 0$, but $d_{12}^2 + d_{13}^2 - d_{23}^2 = 0$ (i.e. $\mathbf{m}_3\cdot\mathbf{m}_1 = 0$, and $\mathbf{m}_3\cdot\mathbf{m}_2 = 0$, but $d_{12}^2 + d_{13}^2 - d_{23}^2 \neq 0$), the solution can be obtained by:

$$\begin{cases} \dfrac{d_{12}^2 + d_{13}^2 - d_{23}^2}{2d_{12}^2}(y_2\mathbf{m}_2 - \mathbf{m}_1)\cdot(y_2\mathbf{m}_2 - \mathbf{m}_1) \\ \quad +\mathbf{m}_1\cdot(y_2\mathbf{m}_2 - \mathbf{m}_1) = 0, \\ y_3^2\mathbf{m}_3\cdot\mathbf{m}_3 = \dfrac{d_{13}^2}{d_{12}^2}(y_2\mathbf{m}_2 - \mathbf{m}_1) \\ \quad \cdot(y_2\mathbf{m}_2 - \mathbf{m}_1) - \mathbf{m}_1\cdot\mathbf{m}_1 \end{cases} \tag{C3}$$

The geometric meaning of the conditions is:

$$\angle\mathbf{M}_1\mathbf{OM}_3 = 90^\circ, \text{ and } \angle\mathbf{M}_2\mathbf{OM}_3 = 90^\circ, \text{ but } \angle\mathbf{M}_2\mathbf{M}_1\mathbf{M}_3 \neq 90^\circ$$

Case (iv).: When $\mathbf{m}_3\cdot(y_2\mathbf{m}_2 - \mathbf{m}_1) = 0$, and $\mathbf{m}_3\cdot\mathbf{m}_2 = 0$, and $d_{12}^2 + d_{13}^2 - d_{23}^2 = 0$ (i.e. $\mathbf{m}_3\cdot\mathbf{m}_1 = 0$, and $\mathbf{m}_3\cdot\mathbf{m}_2 = 0$, and $d_{12}^2 + d_{13}^2 - d_{23}^2 = 0$), by substituting $d_{12}^2 + d_{13}^2 - d_{23}^2 = 0$ into (C3), the solution of (5.5) is given as:

$$\begin{cases} y_2 = \dfrac{\mathbf{m}_1\cdot\mathbf{m}_1}{\mathbf{m}_1\cdot\mathbf{m}_2}, \\ y_3^2\mathbf{m}_3\cdot\mathbf{m}_3 = \dfrac{d_{13}^2}{d_{12}^2}(y_2\mathbf{m}_2 - \mathbf{m}_1) \\ \quad \cdot(y_2\mathbf{m}_2 - \mathbf{m}_1) - \mathbf{m}_1\cdot\mathbf{m}_1 \end{cases} \tag{C4}$$

The geometric conditions that (C4) has a unique group of positive solutions are:

$$\angle\mathbf{M}_1\mathbf{OM}_3 = 90^\circ, \text{ and } \angle\mathbf{M}_2\mathbf{OM}_3 = 90^\circ, \text{ and } \angle\mathbf{M}_2\mathbf{M}_1\mathbf{M}_3 = 90^\circ, \text{ and } \angle\mathbf{M}_1\mathbf{M}_3\mathbf{M}_2 < \angle\mathbf{M}_1\mathbf{OM}_2 < 90^\circ$$

Remark 2. There is always $\mathbf{m}_1\cdot\mathbf{m}_2 \neq 0$ in (C4). This is because the space configuration for $\mathbf{m}_3\cdot\mathbf{m}_1 = 0$, $\mathbf{m}_3\cdot\mathbf{m}_2 = 0$, $d_{12}^2 + d_{13}^2 - d_{23}^2 = 0$, and $\mathbf{m}_1\cdot\mathbf{m}_2 = 0$ is not realizable in 3D Euclidean space.

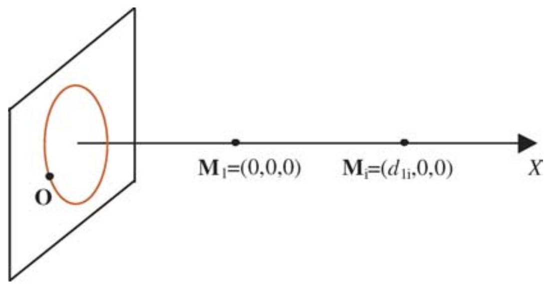

Figure 7. The locus of the optical center determined by collinear control points.

5.2. *Quasi-Analytical Solutions of the PnP Problem when $n \geq 4$*

Like in Section 5.1, the quasi-analytical solutions of the PnP problem with $n > 3$ can be similarly obtained. However, due to the limited space, they will not be reported. The interested reader is referred to [15] for details. The following is an illustrative example:

When $\mathbf{m}_i \cdot (y_2\mathbf{m}_2 - \mathbf{m}_1) \neq 0, i = 3 \ldots n$, (5.4) can be changed as:

$$\begin{cases} y_i = \dfrac{\left(\left(d_{12}^2 + d_{1i}^2 - d_{2i}^2\right) y_2\mathbf{m}_2 + \left(d_{12}^2 - d_{1i}^2 + d_{2i}^2\right)\mathbf{m}_1\right)\cdot(y_2\mathbf{m}_2 - \mathbf{m}_1)}{2d_{12}^2\mathbf{m}_i \cdot (y_2\mathbf{m}_2 - \mathbf{m}_1)}, \quad i = 3\ldots n \\ \dfrac{d_{1i}^2}{d_{12}^2}(y_2\mathbf{m}_2 - \mathbf{m}_1)\cdot(y_2\mathbf{m}_2 - \mathbf{m}_1) = (y_i\mathbf{m}_i - \mathbf{m}_1)\cdot(y_i\mathbf{m}_i - \mathbf{m}_1), \quad i = 3\ldots n \\ \dfrac{d_{1i}^2 + d_{1j}^2 - d_{ij}^2}{2d_{12}^2}(y_2\mathbf{m}_2 - \mathbf{m}_1)\cdot(y_2\mathbf{m}_2 - \mathbf{m}_1) = (y_i\mathbf{m}_i - \mathbf{m}_1)\cdot(y_j\mathbf{m}_j - \mathbf{m}_1), \quad i, j = 3\ldots n, i \neq j \end{cases} \tag{CN1}$$

In (CN1), we have in total $(n-2) + \binom{n-2}{2} = \frac{(n-1)(n-2)}{2}$ univariate quartic equations with respect to y_2. This is the general case discussed in [12].

6. Degenerate Cases for Coplanar or Collinear Control Points

In the previous sections, we have assumed that no three of the control points are collinear, no four of the control points are coplanar. In this section, we consider the degenerate cases where either all control points are coplanar or collinear.

6.1. *Coplanar Case*

If n control points with $n \geq 4$ are all coplanar, then the following determinants are zero:

$$\begin{aligned} &[\mathbf{M}_2 - \mathbf{M}_1,\ \mathbf{M}_3 - \mathbf{M}_1,\ \mathbf{M}_i - \mathbf{M}_1] \\ &\quad = [x_2\mathbf{m}_2 - x_1\mathbf{m}_1, x_3\mathbf{m}_3 - x_1\mathbf{m}_1, x_i\mathbf{m}_i - x_1\mathbf{m}_1] \\ &\quad = 0, \quad i = \cdots n \end{aligned} \tag{6.1}$$

Theorem 5. *When all control points are generically coplanar (no three of them are collinear), then solving $x_i, i = 1\ldots n$, by (3.4) and (6.1) are equivalent to solving ($\boldsymbol{R}, \boldsymbol{t}$) by (3.2) such that $\boldsymbol{R}$ is a rotation matrix.*

(3.4) and (6.1) are nonlinear equations with respect to x_i, while, (3.2) are linear with respect to variables x_i and $\boldsymbol{r}_1$, $\boldsymbol{r}_2$, $\boldsymbol{t}$. When $n \geq 4$, from (3.2), these variables can be linearly solved out up to a scalar. The scalar can be determined by the property $\boldsymbol{r}_1 \cdot \boldsymbol{r}_2 = 1$. So solving the PnP problem for coplanar case should choose (3.2) rather than (3.4) or (2.1). But, the complete solutions of (3.4) and (6.1) can also be obtained easily via the above Theorem 5.

6.2. *Collinear case*

Theorem 6. *When all control points are generically collinear (the optical center is not collinear with them), then the number of solutions ($\boldsymbol{R}, \boldsymbol{t}$) of (3.2) is infinitely large, while the distances X_i of (2.1) from the optical center to control points can be uniquely determined.*

In this case, the configuration of the optical center and control points is shown in Fig. 7. A detailed analysis can be found in [15] and here is a summary: By taking the line through the collinear control points as X-axis, then only $\boldsymbol{t}$ and $\boldsymbol{r}_1$ are uniquely determined. Hence only the first coordinate of the optical center $\boldsymbol{O} = -\boldsymbol{R}^\tau \boldsymbol{t}$ is determined uniquely, and $\boldsymbol{O}$ is therefore determined up to one degree of freedom. In fact the locus of $\boldsymbol{O}$ is a circle on the plane orthogonal to *line through the control points, and its center is the intersection point of the plane with this line.*

7. Conclusions

We clarify the relationships and the upper bounds of the number of solutions for the whole PnP problem under different definitions. In addition, a new and simple elimination method for solving the PnP problem is provided. Some degenerate cases, such as coplanar or collinear control points, are also discussed. In our further work, the optimized algorithm based on the proposed depth equations and its sensitivity will be studied.

Appendix A: Proof of Result 3

With the coordinate system as that in Fig. 1, let $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{t}$ be a group of solutions determined by $\mathbf{M}_i$, $i = 1, 2, 3$, and let $\mathbf{r}_3 = \mathbf{r}_1 \times \mathbf{r}_2$, $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$, $\mathbf{R}' = (\mathbf{r}_1, \mathbf{r}_2, -\mathbf{r}_3)$. Because no four points are coplanar, the third coordinates of $\mathbf{M}_j$, $j = 4\ldots n$ are nonzero.

If both ($\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$, $\mathbf{t}$) and ($\mathbf{R}' = (\mathbf{r}_1, \mathbf{r}_2, -\mathbf{r}_3)$, $\mathbf{t}$) are the solutions, then:

$$\begin{cases} x_j \mathbf{m}_j = \mathbf{R}\mathbf{M}_j + t \\ x'_j \mathbf{m}_j = \mathbf{R}'\mathbf{M}_j + t \end{cases} \qquad j = 4\ldots n.$$

The first equation subtracts the second equation yields:

$$(x_j - x'_j)\mathbf{m}_j = 2c_j \mathbf{r}_3, \quad \text{or} \quad \mathbf{m}_j \approx \mathbf{r}_3 \qquad \text{(A)}$$

where c_j are the third coordinates of $\mathbf{M}_j$, so it is nonzero. Then, the right side of (A) is nonzero and it follows that the left side is nonzero too. On the other hand, there is $(\mathbf{Rt})(0010)^\tau = \mathbf{r}_3$, i.e. $\mathbf{r}_3$ is the image of the point at infinity of Z-axis. So from (A), $\mathbf{m}_j$ is also the image of the point at infinity of Z-axis. It is equivalent to that the line through the optical centers and $\mathbf{M}_j$ should be parallel to Z-axis. Thus, by the used coordinate system, there are $\mathbf{OM}_j \perp \mathbf{M}_1\mathbf{M}_2\mathbf{M}_3$, $\mathbf{O}'\mathbf{M}_j \perp \mathbf{M}_1\mathbf{M}_2\mathbf{M}_3$, where $\mathbf{O} = -\mathbf{R}^\tau\mathbf{t}$, $\mathbf{O}' = -\mathbf{R}'^\tau\mathbf{t}$ are the optical centers symmetric with respect to $\mathbf{M}_1\mathbf{M}_2\mathbf{M}_3$.

If $(\mathbf{R}_1, \mathbf{R}'_1)$ is another pair of solutions, then the corresponding optical centers $\mathbf{O}_1 = -\mathbf{R}_1^\tau\mathbf{t}_1$, $\mathbf{O}'_1 = -\mathbf{R}_1^{\tau\prime}\mathbf{t}_1$ also satisfy $\mathbf{O}_1\mathbf{M}_j \perp \mathbf{M}_1\mathbf{M}_2\mathbf{M}_3$, $\mathbf{O}'_1\mathbf{M}_j \perp \mathbf{M}_1\mathbf{M}_2\mathbf{M}_3$ and are symmetric with respect to $\mathbf{M}_1\mathbf{M}_2\mathbf{M}_3$. Therefore $\mathbf{O}$, $\mathbf{O}'$ $\mathbf{O}_1$, $\mathbf{O}'_1$, and $\mathbf{M}_j$ are collinear, see Fig. 8.

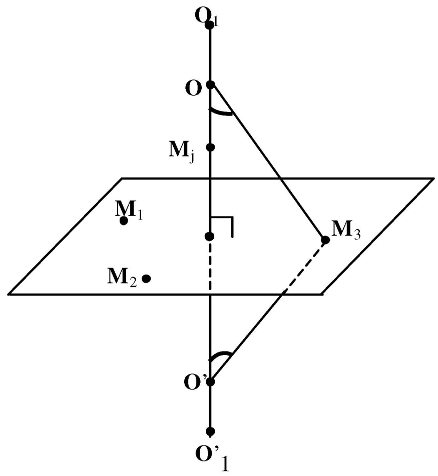

Figure 8. $\mathbf{O}$ and $\mathbf{O}'$, $\mathbf{O}_1$ and $\mathbf{O}'_1$ are two pairs of symmetric optical centers with respect to $\mathbf{M}_1\mathbf{M}_2\mathbf{M}_3$. $\mathbf{O}$, $\mathbf{O}'$, $\mathbf{O}_1$, $\mathbf{O}'_1$ and $\mathbf{M}_j$ are collinear. If $\angle\mathbf{M}_i\mathbf{O}\mathbf{M}_j = \angle\mathbf{M}_i\mathbf{O}_1\mathbf{M}_j$, then $\mathbf{O}_1$ must be $\mathbf{O}$ or $\mathbf{O}'$.

Moreover, by Definition 1, the angles to every pair of control points from the perspective center are prior known, there should be $\angle\mathbf{M}_i\mathbf{O}\mathbf{M}_j = \angle\mathbf{M}_i\mathbf{O}_1\mathbf{M}_j$ $\angle\mathbf{M}_i\mathbf{O}'_1\mathbf{M}_j$, $j = 4\ldots n$, $i \neq j$, $i = 1\ldots n$. But these equalities are not possible unless $\{\mathbf{O}, \mathbf{O}'\} = \{\mathbf{O}_1, \mathbf{O}'_1\}$. If $\{\mathbf{O}, \mathbf{O}\}' = \{\mathbf{O}_1, \mathbf{O}'_1\}$, we assume $\mathbf{O} = \mathbf{O}_1$, $\mathbf{O}' = \mathbf{O}'_1$, then $\mathbf{R}^\tau t = \mathbf{R}_1^\tau\mathbf{t}_1$, so $\mathbf{R}_3^\tau\mathbf{t} = \mathbf{R}_{31}^\tau\mathbf{t}_1$ where $\mathbf{r}_3$, $\mathbf{r}_{31}$ are the third columns of $\mathbf{R}$, $\mathbf{R}_1$. By (A), we have $\mathbf{m}_j \approx \mathbf{r}_3$, $\mathbf{m}_j \approx \mathbf{r}_{31}$, further by the fact that $\mathbf{r}_3$, $\mathbf{r}_{31}$ are of unitary norm, $\mathbf{r}_3 = \mathbf{r}_{31}$ is followed. By the result and $\mathbf{R}_3^\tau\mathbf{t} = \mathbf{R}_{31}^\tau\mathbf{t}_1$, we have $\mathbf{t} = \mathbf{t}_1$. Further by $\mathbf{R}^\tau\mathbf{t} = \mathbf{R}_1^\tau\mathbf{t}_1$, $\mathbf{R} = \mathbf{R}_1^\tau$ is obtained. Then, $\mathbf{R}' = \mathbf{R}_1^{\tau\prime}$ is obtained too. Thus the solutions $\mathbf{t}$, $\mathbf{R}$ and $\mathbf{R}'$ are identical with $\mathbf{t}_1$, $\mathbf{R}_1$ and $\mathbf{R}'_1$. If $\mathbf{O} = \mathbf{O}'_1$, $\mathbf{O}' = \mathbf{O}_1$, the same result can be inferred. Therefore, there is at most one pair of $(\mathbf{R}_i, \mathbf{R}'_i)$ that can be the solutions.

Also, we can see that (i) of Result 3 is true from the above derivation if there is such a pair as the solutions.

When $n \geq 5$, if there is such a pair as the solutions, then by (A) we have $\mathbf{m}_4 \approx \mathbf{m}_j \approx \mathbf{r}_3$, $j = 5\ldots n$. So $\mathbf{M}_4\mathbf{M}_j \perp \mathbf{M}_1\mathbf{M}_2\mathbf{M}_3$. Also since $\mathbf{OM}_j \perp \mathbf{M}_1\mathbf{M}_2\mathbf{M}_3$, $\mathbf{O}'\mathbf{M}_j \perp \mathbf{M}_1\mathbf{M}_2\mathbf{M}_3$, $j = 4\ldots n$, the points $\mathbf{O}$, $\mathbf{O}'$ and $\mathbf{M}_j$ with $j = 4\ldots n$ are all collinear. Thus, $\angle\mathbf{M}_4\mathbf{O}\mathbf{M}_5 = \angle\mathbf{M}_4\mathbf{O}'\mathbf{M}_5 = 0$. It follows that by the angles to every pair of control points from the perspective center are prior known, the possible optical center $\mathbf{O}_x$ must be on the line $\mathbf{OO}'$ in order to be $\angle\mathbf{M}_4\mathbf{O}_x\mathbf{M}_5 = 0$. Moreover, by the proof that there is at most one pair $\mathbf{O}$, $\mathbf{O}'$ that can be the solutions (the paragraph before Fig. 8), we have the conclusion that there are only the two solutions ($\mathbf{R}$, $\mathbf{t}$) and ($\mathbf{R}'$, $\mathbf{t}$). Thus, (ii) of Result 3 is also proved.

Acknowledgments

This work was supported by the National Key Basic Research and Development Program (973) under grant No. 2002CB312104, and National Natural Science Foundation of China under grant No. 60033010, 60303021.

References

1. M.A. Abidi and T. Chandra, "A new efficient and direct solution for pose estimation using quadrangular targets: Algorithm and evaluation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 17, No. 5, pp. 534–538, 1995.
2. A. Ansar and K. Daniilidis, "Linear pose estimation from points or lines," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 25, No. 5, pp. 578–589, 2003.
3. D. Dementhon and L.S. Davis, "Exact and approximate solutions of the perspective-three-point problem," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 14, No. 11, pp. 1100–1105, 1992.
4. M.A. Fischler and R.C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, Vol. 24, No. 6, pp. 381–395, 1981.
5. P.D. Fiore, "Efficient linear solution of exterior orientation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 2, pp. 140–148, 2001.
6. X.S. Gao, X.R. Hou, J.L. Tang and H.F. Cheng, "Complete solution classification for the perspective-three-point problem," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 25, No. 8, pp. 930–943, 2003.
7. J.A. Grunert, "Das pothenotische problem in erweiterter gestalt nebst bber seine anwendungen in der geodäsie," *Grunerts Archiv für Mathematik und Physik*, Band 1, pp. 238–248, 1841.
8. R.M. Haralick, C. Lee, K. Ottenberg and M. Nölle, "Analysis and solutions of the three point perspective pose estimation problem," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 1991, pp. 592–598.
9. R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2000.
10. R. Horaud, B. Conio, and O. Leboulleux, "An analytic solution for the perspective 4-point problem," *Computer Vision, Graphics, Image Processing*, Vol. 47, No. 1, pp. 33–44, 1989.
11. Z.Y. Hu and F.C. Wu, "A note on the number of solutions of the non-coplanar P4P problem," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 4, pp. 550–555, 2002.
12. L. Quan and Z.D. Lan, "Linear *n*-point camera pose determination," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 21, No. 8, pp. 774–780, 1999.
13. C. Su, Y.Q. Xu, H. Li, and S.Q. Liu, "Necessary and sufficient condition of positive root number of perspective-three-point problem," *Chinese Journal of Computers* (in Chinese), Vol. 21, No. 12, pp. 1084–1095, 1998.
14. W.J. Wolfe, D. Mathis, C.W. Sklair and M. Magee, "The perspective view of three points," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 13, No. 1, pp. 66–73, 1991.
15. Y.H. Wu and Z.Y. Hu, "PnP problem revisited," Technical Report, RV-NLPR, Institute of Automation, Chinese Academy of Sciences, 2003.

Yihong Wu received her Bachelor of Science degree in Mathematics from Shanxi Yanbei Normal College in 1995; a Master of Science degree in Computational Algebra from Shaanxi Normal University in 1998; a Doctor of Science degree in Geometric Invariants and Applications from MMRC, Institute of Systems Science, Chinese Academy of Sciences, in 2001. From June 2001 to July 2003, she did her postdoctoral research in NLPR, Institute of Automation, Chinese Academy of Sciences. After then, she joined NLPR as an associate professor. Her research interests include polynomial elimination and applications, geometric invariant and applications, automated geometric theorem proving, camera calibration, camera pose determination, and 3D reconstruction. She has published more than 15 papers on major international journals and major international conferences.

Zhanyi Hu was born in Shanxi province, P. R. China in 1961. He received the B.S. Degree in Automation from the North China University of Technology in 1985, the Ph.D. Degree (Docteur d'Etat) in Computer Science from the University of Liege, Belgium, in Jan. 1993. Since 1993, he has been with the Institute of Automation, Chinese Academy of Sciences. From May 1997 to May 1998, he also acted as a visiting scholar of Chinese University of Hong Kong on invitation. Dr. Hu now is a Research Professor of Computer Vision, a member of the Executive Expert Committee of the Chinese National High Technology R&D Program, a deputy editor-in-chief for Chinese Journal of CAD and CG, and an associate editor for Journal of Computer Science and Technology. His current research interests include Camera Calibration, 3D Reconstruction, Feature Extraction, Vision Guided Robot Navigation etc. Dr. Hu has published more than 70 peer-reviewed papers on major national and international journals.