Deeper Understanding on Solution Ambiguity in Estimating 3D Motion Parameters by Homography Decomposition and its Improvement

A dissertation presented for the degree of Doctor of Engineering

He Liu

Supervisor: Hiroyuki Hase
Department of Fiber Amenity Engineering
Graduate School of Engineering
University of Fukui
Japan

March, 2012
CONTENTS

Acknowledgements........................................................................................................ iii

Abstract ....................................................................................................................... iv

Chapter 1   Introduction ..................................................................................1

Chapter 2   Homography Mapping and Homography Decomposition for 3D Motion Parameters Estimation.................................................................10
  2.1 Homography Mapping...................................................................................10
  2.2 Faugeras’ Homography Decomposition Method for 3D Motion Parameters Estimation .................................................................18
  2.3 Zhang’s Homography Decomposition Method for 3D Motion Parameters Estimation .................................................................20
  2.4 Iterative Method to Estimate the Real Rotation Center Coordinate on the Object ................................................................................26
  2.5 Solution Ambiguity Problem of 3D Motion Parameters Estimation by Homography Decomposition..................................................30

Chapter 3   Two Dependences of Solution Ambiguity by Homography Decomposition..........................................................................................35
  3.1 Dependence on 3D Motion Parameters of Solution Ambiguity by Homography Decomposition.................................................................37
  3.1.1 Dependence on 3D Motion Parameters and its Theoretical Proof ........................................................................................................37
3.1.2 Geometrical Proof of Dependence on 3D Motion Parameters
..............................................................................................................60

3.2 Dependence on Object Size of Solution Ambiguity by Homography Decomposition.......................................................63

Chapter 4  Constraint Condition for Homography Decomposition .....69

Chapter 5  Simulation Experiment and Results.................................73

  5.1 The Convergence of the Iterative Rotation Center Estimating Method .........................................................................................73

  5.2 The Two Dependences of Solution Ambiguity Problem in Estimating 3D Motion Parameters by Homography Decomposition Method............................................................................76

  5.3 Constraint Condition of Joined Two Planes .................................83

Chapter 6  Conclusion..............................................................................87

References.................................................................................................89

Publication List .........................................................................................93
Acknowledgements

First of all, I would like to express my sincere gratitude to my academic supervisor, Professor Hiroyuki Hase (Faculty of Engineering, University of Fukui, Japan). He gave me the second chance to finish my Ph.D. research. From him, I have acquired not only the continuous guidance and the instructive suggestion on academy, but also the fabulous direction on my life.

Highly tribute shall be paid to another my academic supervisor, Associate Professor Shogo Tokai (Faculty of Engineering, University of Fukui, Japan), whose constructive advice and useful suggestion are one of the decisive factors to my research. I am deeply grateful of his help in the completion of this thesis.

I am also deeply indebted to all the other tutors and teachers in Faculty of Engineering in University of Fukui for their direct and indirect help to me.

Finally, I am indebted to my parents for their never-ending love and encouragement.
Abstract

3D pose estimation or 3D motion estimation is one of the most important problems in the computer vision field. For this problem, a lot of methods for fixing the shape restoration problem from consecutive images have been proposed up until now, including non-linear methods and linear methods.

When the object is a plane, the camera displacement parameters can be extracted (assuming that the intrinsic camera parameters are known) from the homography matrix that can be measured from two views. This process is called homography decomposition. The standard algorithms for homography decomposition give numerical solutions using the singular value decomposition (SVD) of the matrix, proposed by Faugeras and Zhang.

There are generally at least two possible solutions only under the visual reality constraint by using only two pieces of images. The ambiguity of the two solutions cannot be resolved if there is no further a priori information available. This is the solution ambiguity problem in estimating the 3D motion parameters by homography decomposition. From then on, the inevitability of the ambiguity problem to estimate solutions has been widely accepted for nearly 30 years. It is considered to be one of the basic properties and mathematical defects in homography decomposition, even today.

My research focuses on this solution ambiguity problem of homography decomposition. The different results have been found in my research: solution ambiguity is not inevitable, and the unique solution can be obtained conditionally. Two kinds of dependences of the solution ambiguity problem in estimating the 3D motion parameters by homography decomposition have been clarified for the first time:

(1) My research pointed out the dependence of solution ambiguity on 3D motion parameters for the fixed object size, and defined a criterion about the 3D motion parameters to obtain the unique estimated solution: “If all distances between the
feature points and the camera do not become closer after 3D motion, the solution ambiguity problem can be avoided”. This conclusion has been geometrically and theoretically derived and it is the first time this dependence has been mentioned. However, because the constraint for the theoretical proof is more general than that in the previous work, this theory should be more reliable.

(2) My research pointed out the dependence of solution ambiguity on object size for the fixed 3D motion parameters and explained it with geometry. Meanwhile, a criterion about the object size to obtain the unique estimated solution has been defined: “If the feature region is large enough, or the range of 3D movements is relatively small compared to the size of the feature region, the solution ambiguity problem can be avoided by using homography decomposition”. This dependence and its geometrical proof have been presented for the first time until now.

Besides the two kinds of dependences of the solution ambiguity problem in estimating the 3D motion parameters by homography decomposition, my research also proposes a kind of constraint condition of joined two planes, in order to guarantee the unique real estimated solution. The effectiveness of the constraint condition has been tested by the simulation experiments. The occurrence ratio of the ambiguous solutions is smaller than 0.3‰.

My research clarified the mechanism for solution ambiguity obtained by homography decomposition in the estimation of 3D motion parameters. Both dependencies on 3D motion parameters and object size have been theoretically and experimentally verified. All of the theories and the constraint condition have been verified by simulation experiments used the huge database.

In my research, I do not have any intention of proposing a new approach to solving the solution ambiguity problem of homography decomposition more effectively in my research. The only purpose of my research is to give an accurate interpretation and a complete description of facts that have been misunderstood for a long time. Therefore, the simulation results should have been sufficiently accurate, so that a demonstration with real images was not necessary in my research.
Chapter 1  Introduction

Estimation of object 3D motion parameters or 3D reconstruction of camera positions and scene structure based on features from multiple viewpoints images has been studied for over several decades. Several methods for vision-based robot control need an estimation of the camera displacement (i.e. rotation and translation) between two views of an object [1, 2, 3]. Techniques for 3D motion estimation can be applied to many fields, such as 3D reconstruction, image processing, robot controlling, and so on. It is also an important part of aerial image interpretation, object recognition, etc.

3D reconstruction is an important problem and still remains very difficult. Over the years, this problem has been related to structure from motion, stereo vision, pose determination, etc. And it has been addressed by many researchers [4, 5]. Recently, projective geometry has been used to perform 3D reconstruction (based on uncalibrated cameras) up to an unknown projective model [7, 8, 9, 10].

For this problem, a lot of methods (including the factorization method proposed by Tomasi and Kanade [11]) for fixing the shape restoration problem from consecutive images have been proposed up until now. The factorization method is based on an orthogonal projection. It can restore relative 3D shape and 3D motion at the same time from consecutive images. Aside from this, some improvements like the perspective projection model have been done. In these methods, corresponding features observed from a lot of consecutive frames are necessary for calculation. However, some lack of feature points can not be avoided due to occlusion.

Meanwhile, the nonlinear method is another option for this problem. For at least three or four non-collinear corresponding points for the two different 3D poses, exact estimation solutions can be computed: A fourth- or fifth-degree polynomial system can be formulated using geometrical invariants of the observed points and the problem can be solved by finding solutions of the polynomial system [12, 13]. Many methods are proposed because the solutions are hardly to get deterministically. One of the most typical methods is Gauss-Newton iterative method [14]. The Gauss-Newton
iterative method is used to solve the nonlinear least-squares formulations. And Levenberg-Marquardt Method is also a useful one which is a sort of advanced Gauss-Newton method. This is one of the most reliable optimization methods currently in use. But as a kind of iterative method, both of these two kinds of methods rely on a good initial guess to guarantee to converge to the correct solution. And the 3D shape of the object is necessary.

Then, Lu and Hager etc. proposed Orthogonal Iteration (OI) Method [15], which is different from the traditional methods which formulate pose estimation based on the estimation error minimization in the image space; this method is based on the estimation error minimization in the object space. The computation speed of OI method is relatively faster than the traditional methods. Although it can guarantee the global convergence for this method, it is hardly to prove that OI method will converge to the correct solutions. And in OI method, the 3D shape of the object is necessary.

However, for all of the iterative methods, there always contains a problem which the estimation process will be trapped into the local minima unless a good initial guess is available. And most of this kind of methods are based on the situation which the 3D shaped of the object is known. Therefore, Liu, Hase and Tokai have proposed a kind of non-linear method to obtain the 3D motion parameters by LSM (Least Squares Method) based on a more convenient perspective projection model [16, 17]. Because the derived equations are hardly solved deterministically, the steepest descent method (SDM) is used to find the solution. However, this iterative procedure sometimes happens to fall into a local minimum in the case of a few feature points depending on the initial values. Then this problem is resolved by using the annealing algorithm. Although both the signs and the value of the solutions can be determined automatically in this method, the computation cost is very huge and it needs too much time for a complete estimation.

Specifically, when the object is a plane, it is common to address the problem of 3D motion parameters estimation by reconstructing a 3D scene based on externally uncalibrated cameras, while assuming internal calibration is known. The problem is: given a correspondence of two sets of coplanar points from two unknown positions
before and after an arbitrary 3D motion, 3D points in Euclidean space can be reconstructed. Therefore, in common situations, 3D motion parameters estimation means solving method for the relative geometry before and after the 3D movement between from two pieces of images taken by an uncalibrated camera, and the 3D plane that gives rise to the coplanar correspondences in the two images.

Faugeras [6] and Hartley et al [4, 7] are two of the earliest efforts to address the problem of 3D reconstruction based on two uncalibrated images. Faugeras showed that given five non-coplanar correspondences and epipoles, or given the fundamental matrix, 3D structures can be reconstructed up to a collineation of projective transforms. This result was then extended to the affine case, where he showed that given four non-coplanar points and epipoles, 3D structures can be recovered under an affine transform with three unknown parameters.

Hartley et al independently reached the same conclusion by using matrix theory to linearly decompose the essential matrix. Previous research results [18, 19] showed that if the internal calibration of cameras in known, then it is possible to determine the relative motion (or geometry) between the two cameras and the relative locations of the 3D points corresponding to the matched points in the two views from the essential matrix, which needs at least eight correspondences. Hartley then showed that this is also true even when the focal lengths of the two cameras are unknown, and Hartley et al went on to show that if the internal calibration of cameras is completely unknown, it is not possible to recover the relative geometry and locations of 3D points unambiguously in a Euclidean space. In this case, the recovered 3D locations of the points and the camera geometry are subject to a $4 \times 4$ projective transform matrix, and absolute Euclidean coordinates of the 3D points can be computed only when a set of ground control points are known. Later Mohr et al [8] solve the same problem by making full use of the redundancy in multiple images to directly solve for a global least mean square solution.

Shashua [9] explored a new projective invariant, which he referred to as projective depth. Using this invariant, he showed that given four non-coplanar correspondences and epipoles of two views, 3D reconstruction can be achieved under
either a projective transformation, or an affine transformation, depending on how the views were produced. As compared with previous work, the main contribution of this work (other than the exploration of the projective depth invariant) is that orthographic and perspective projections are treated in the same manner, and the computation does not need to recover the camera transformation first (i.e. structure without motion).

Ponce et al [20] discussed several different cases under a projective transform and proposed algorithms to reconstruct 3D structures. Some of them assume weak calibration, i.e. known epipoles, while others do not. More recently, Shashua and Navab [21] presented a novel theory called “relative affine structure”. Based on this theory, they developed a unified method for recovering 3D structures. Again, at a minimum, two views, four non-coplanar points, and the epipoles are required. The 3D structure is obtained under an affine transform with three unknown parameters. Hartley [22] also extended his previous reconstruction method based on points to a new algorithm using lines. Still, this reconstruction is under projective space, and it assumes at least three views. Sawhney [10] attacked the same problem slightly differently. Instead of aiming at point-based 3D structure reconstruction, he used planar motion parallax and image warping techniques to present a unified framework for intrinsic 3D shape reconstruction for three projection models: weak, para, and full perspective. A similar treatment using a different approach was also independently done by Kumar and Anandan [23] using motion parallax. Recently, Hartley [24, 25] proposed to use a trifocal tensor to reconstruct 3D scene up to projectivity from three uncalibrated views. This is a linear algorithm, and is a unified approach in the sense that it can apply either to lines or to points (or to the combination of lines and points). In particular, he showed that this trifocal tensor is essentially identical to the set of coefficients introduced by Shashua [26] to effecting point transfer in the three view case. Another major contribution of this work is that Hartley showed that the minimum requirement for 3D projective reconstruction from 3 views with the same camera is either 7 point correspondences, or 13 line correspondences. In the point case, the minimum number 7 is consistent with the results of Maybank et al [27] and Faugeras [28], except that in their work, a Euclidean reconstruction was possible
because they explicitly recovered the internal camera parameters. However, they
needed at least three motions, and assumed that the same camera was used to acquire
all views. In the line cases, the minimum number 13 is consistent with Weng et al [29],
except that their work assumed that the cameras were calibrated.

In summary, it has been proven that in the case of two uncalibrated views, it is
impossible to recover the 3D scene in a Euclidean space. The best one which can done
is to recover the 3D scene up to an arbitrary projectivity, and to do so require at least
seven correspondences. If three views are available, and assuming the same camera is
used for all three views, then it is possible to get scaled Euclidean 3D reconstruction
[30, 31].

When the object is a plane, the camera displacement can be extracted from the
homography matrix that can be measured from two views. This process is called
homography decomposition. The standard algorithms for homography decomposition
give numerical solutions using the singular value decomposition (SVD) of the matrix,
proposed by Faugeras [32] and Zhang [33].

In Faugeras’ paper [32], he showed the fact that the environment is piecewise
linear provides a powerful constraint on the kind of matches that exist between two
images of the scene when the camera (or the object) motion in unknown. For points
and lines located in the same plane, the correspondence between the two positions (or
the two cameras) is a collineation. He showed that the unknowns (the camera motion,
or the object motion, and the plane equation) can be recovered, in general, from an
estimate of the matrix of this collineation, homography matrix. The twofold
ambiguity that remains can be removed by looking at a second plane, by taking a third
view of the same plane, or by using a priori knowledge about the geometry of the
plane considered. He then showed how to combine the estimation of the matrix of
collineation and the acquirement of point and line matches between the two images,
by a strategy of Hypothesis Prediction and Testing guided by Kalman Filter. He
finally showed how his approach can be used to calibrate a system of cameras.

His work extended what Tsai and Huang [34] proposed in their researches. Like
all the previous authors, he assumed that, in the motion case, only one motion is
present; in his case, even though it is not essential in the mathematical proofs, he assumed for simplicity that the camera is moving in an otherwise static environment. He made full use of the fundamental property that if two cameras are forming the image of a plane, then the correspondence between the two retinas has a very simple analytical form: it is a collineation, i.e. it is linear in projective coordinates. This implies in particular that, because of the fundamental duality property of projective geometry, points and lines play exactly the same role as tokens. He characterize (like Tsai and Huang [34]) the relationship between this collinear and the geometry of the problem. He then proved (in a somewhat simpler way than Tsai and Huang) that, given the matrix of the collineation, the position and orientations of the second camera (or the object position after 3D motion) and the plane with respect to the first one (or the object position before 3D motion) can be recovered. He analyzed in detail the number of solutions and provided a geometric interpretation of the degenerate cases.

Zhang’s Method [33] is more advanced than Faugeras and Lustman’s [32]. In his paper, he presented a method that recovers the 3D structures in a Euclidean space. This is an extension of the work by Zhang and Hanson [35]. Like the previous work, at a minimum his method also needs two views and four correspondences. Unlike the previous work, which assumes four non-coplanar correspondences, Zhang assumes four coplanar correspondences that are not collinear. Moreover, the 3D scene structures are recovered in Euclidean space up to two solutions with a uniform scaling factor, as opposed to a family of solutions in a projective space; this scaling parameter has an explicit physical meaning which is the distance of the first camera center from the 3D plane formed by the four points given in the correspondences. If this distance is known a priori, then a complete 3D Euclidean reconstruction can be obtained up to two solutions. Without any a priori knowledge, it is shown that these two solutions are indistinguishable. The other difference from the previous work in 3D reconstruction is that there Zhang’s method assumed that the internal calibration is known, as opposed to assuming weak calibration, or completely unknown calibration. The basic idea is that he first found the homography mapping between the two cameras (or two object
positions before and after 3D motion) using the four given correspondences. The homography matrix is decomposed to obtain the relative pose between the two cameras (or the motion parameters of the object). Finally, 3D structures are reconstructed based on the recovered relative pose. He showed that this assumption was necessary for 3D reconstruction in Euclidean space in his paper.

Note that the method proposed by Zhang is different from early work on 3D reconstruction based on the essential matrix [18, 19]. The differences are reflected in two ways. First, the homography matrix is not the same as the essential matrix. Second, using the essential matrix gives only one constraint for each correspondence. Thus the minimum number of correspondences required by using the essential matrix is 8 if linear constraints are used; or 7 if nonlinear constraints are used. However, with the homography matrix, each correspondence produces two equations, which reduces the minimum number of required correspondences to 4.

Faugeras and Lustman [32] and Maybank [27] showed that a homographic transformation between two cameras (or two object positions before and after 3D motion) can be decomposed, and in general, there are two real solutions. Weng et al [36] also came to the same conclusion. However, they only applied the decomposition of the homography matrix to planar 3D reconstruction. In Zhang’s paper, he proposed a different way to decompose this matrix and perform a case by case analysis of different geometric situations. Finally, different analytical closed-form solutions were developed based on all the possible cases. This enabled them to reconstruct any 3D scene, and owing to this closed-form solution, the computation is very inexpensive and fast. It was also shown that the proposed algorithm is optimal.

In brief, whatever Faugeras and Lustman’s Method [32], or Zhang’s method [33], the standard algorithm for homography decomposition obtains numerical solutions using the singular value decomposition of the matrix. It is shown that in the general case there are two possible solutions by the homography decomposition. This numerical decomposition has been sufficient for many computer and robot vision applications. However, when dealing with robot control applications, an analytical procedure to solve the decomposition problem would be preferable (i.e. analytical
expressions for the computation of the camera displacement directly in terms of the components of the homography matrix). Indeed, the analytical decomposition allows us the analytical study of the variations of the estimated camera pose in the presence of camera calibration errors (or the 3D motion parameters before and after object movement).

Therefore, Malis and Vargas [37] proposed a totally new method, by which the insights on the robustness of vision-based control laws can be obtained. The new method provided analytical expressions for the solutions of the problem, instead of the traditional numerical procedures. The main advantage of this method is that it provided a deeper understanding on the homography decomposition problem. For instance, it allowed to obtaining the relations among the possible solutions of the problem. Thus, new vision-based robot control laws can be designed. For example, the control schemes proposed in Mail and Vargas’ paper combine the two final solutions of the problem (only one of them being the true one) assuming that there is no a priori knowledge for discerning among them.

However, the main disadvantage of Malis and Vargas’ method is that the theoretical situation is too complex to be used in the common 3D motion parameters estimation. There are nearly 10 constraints for this method need to be satisfied, so it is hardly to fix every situation of the real 3D reconstruction in each application. Therefore, Faugeras and Lustman’s Method [32], or Zhang’s method [33] still will be the best options for the 3D motion parameters estimation or the 3D reconstruction.

On the other hand, the solution ambiguity problem becomes the main disadvantage of the traditional homography decomposition methods. There are generally at least two possible solutions only under the visual reality constraint by using only two pieces of images. Zhang concluded that “…this ambiguity of the two solutions cannot be resolved if there is no further a priori information available…” [38] on this solution ambiguity problem of homography decomposition. From then on, the inevitability of the ambiguity problem to estimate solutions has been widely accepted for nearly 30 years. It is considered to be one of the basic properties and mathematical defects in homography decomposition, even today.
However, the different results have been found in my research; Solution ambiguity is not inevitable. The appearance of the solution ambiguity problem depends on two factors: if the object size is fixed, it depends on changing the 3D motion parameters; if the 3D motion parameters are fixed, it depends on changing the size of the object.

Therefore, one purpose of my research is to give an accurate interpretation of facts which has been misunderstood for a long time. Meanwhile, I also intent to propose a new improved method using two planes’ constraint to resolve the solution ambiguity problem in estimating 3D motion parameters from only two pieces of images by homography decomposition. The true solution can be gained by this constraint in any conditions of object size and motion parameters.

This research assumes that only two pieces of images are available without any priori constraints. Only the situation in which a unique solution or ambiguous solutions occur under theoretical considerations will be discussed in my research. The theory and the constraint have been verified through simulation experiments which are based on a huge database.

In the following chapters, the basic knowledge of homography mapping and two kinds of homography decomposition methods will be introduced in Chapter 2; in Chapter 3, the theory will be explained in detail; in Chapter 4, an extra constraint will be proposed; All of experiment results of simulation will be introduce in Chapter 5; Finally, the main conclusion of my research will be described in Chapter 6.
Chapter 2  Homography Mapping and Homography Decomposition for 3D Motion Parameters Estimation

In this chapter, I will briefly review the basic theories about homography mapping and homography decomposition in estimating the 3D motion parameters. Two kinds of homography expressions under the two kinds of projection model will be introduced (two cameras model and moving object model). Especially, in the moving object projection model, which is focused on in my research, an original method to estimate the rotation center on the object will also be introduced as the basic theory.

2.1 Homography Mapping

I will totally introduce two kinds of homography matrix expression under two kinds of perspective projection model: one is two cameras model as illustrated in Fig. 1; the other one is the moving object model as illustrated in Fig. 2. The first one is common in 3D reconstruction, and the second one is always be used in the object 3D motion estimation.

First, I will introduce the two cameras perspective projection model [33]. Fig. 1 illustrates geometrical relationship of this situation. Let us assume that the camera at the second position, $O_2$, is first rotated by a rotation matrix $R$ on the position $O_1$, and then translated by a translation vector $t$ to the position $O_2$. Let us assume that both cameras take images of the 3D plane $\pi$, which has a normal vector $n$. Furthermore, assume that all coordinates are represented in the first camera’s coordinate system.
Fig. 1 Geometry of a homography mapping under the two-camera’s perspective projection model
Homogeneous coordinates are used to represent the 2D point \( p \) in the image plane for any 3D point \( P = (X, Y, Z) \). Its corresponding image point in the first camera coordinate system is:

\[
P = \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}, \quad p = \omega \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}
\]

(2-1)

where \( \omega \) is the standard scale factor. The above equations are under the constraints:

\[
x = \frac{X}{Z}, \quad y = \frac{Y}{Z}
\]

(2-2)

There is a similar relationship between the 3D point \( P \) and its projection point \( p' \) in the second camera coordinate system. Now, let \( d \) be the perpendicular distance from the projection center of the first camera to the plane \( \pi \). It can be shown that if \( P \) is on the plane \( \pi \), i.e. \( P \) satisfies:

\[
P \cdot n = d
\]

(2-3)

Furthermore, from Eq. (2-3), Eq. (2-4) can be obtained easily:

\[
1 = \frac{P \cdot n}{d}
\]

(2-4)

Meanwhile, the 3D coordinates \( P \) and \( P' \) of one point in the two camera coordinate systems can be related by the 3D movement expression:

\[
P' = RP + t
\]

(2-5)

Substitute Eq. (2-4) into Eq. (2-5), the following expression can be obtained:

\[
P' = RP + t \cdot \frac{P \cdot n}{d}
\]

\[
= RP + t \frac{n' P}{d}
\]

\[
P' = \begin{pmatrix} R + \frac{t n'}{d} \end{pmatrix} P
\]

(2-6)

Considering the two corresponding projection points \( p \) and \( p' \) of the 3D point coordinates by Eq. (2-1) and (2-2), there is a homographic mapping relationship between \( p \) and \( p' \):
\[ \alpha p' = Hp \]

(2-7)

where \( \alpha \) is a scaling factor accounting for the fact that the representation of \( p \) and \( p' \) are expressed in homogeneous coordinates. \( H \) is the homography mapping matrix, which is a 3×3 matrix and can be shown to be:

\[ H = R + \frac{tn'}{d} \]

(2-8)

Second, I will introduce the homography mapping matrix expression under the moving object perspective projection model. Actually, there are two kinds of situations of the moving object perspective projection model: one is the object is rotated around the camera which is the rotation center; the other one is the object is rotated around the geometry center on itself which is the rotation center. The former is extremely the same as the two cameras perspective projection model, which have already been introduced above. The latter is the major situation in estimating the 3D motion parameters of a moving object. Therefore, it becomes the focus of my research.

Fig. 2 illustrates geometry relationship of the moving object perspective projection model. There is only one static camera \( O \) as the view point. Two 3D feature planes, \( \Pi_1 \) and \( \Pi_2 \), are the same planar object from two arbitrary positions, which are related by a translation vector \( t \) and a rotation matrix \( R \) (around the rotation center \( P_c \)). The distance from the center of the camera \( O \) to the plane \( \Pi_1 \) is \( d \). Let \( P \) be an arbitrary 3D feature point on the plane \( \Pi_1 \), represented in the coordinate system of camera \( O \), and \( Q \) be the corresponding feature point on plane \( \Pi_2 \) in the same coordinate system.

Let \( p \) and \( q \) be the two corresponding feature vectors on the Image Plane respectively. \( P \) and \( Q \) are related by a transformation as Eq. (2-9):

\[ Q = R(P - P_c) + P_c + t \]

(2-9)

The equation for the plane \( \Pi_1 \) is known:

\[ P \cdot n = n'P = d \, . \]

(2-10)
Here, \( \mathbf{n} \) is the normal vector of \( \Pi_1 \).
Fig. 2 Geometry of a homography mapping under the moving object perspective projection model
Substituting Eq. (2-10) for Eq. (2-9), there is

\[ Q = \left[ R + \frac{(I - R)P_c + t}{d} n' \right] P. \]  

(2-11)

According to the rule of perspective projection: \( p = \frac{P}{P_z} f, \quad q = \frac{Q}{Q_z} f, \) Eq. (2-11) can be changed into:

\[ \alpha q = \left[ R + \frac{(I - R)P_c + t}{d} n' \right] p. \]  

(2-12)

Here, \( P_z \) is the z coordinate of \( P \), \( Q_z \) is the z coordinate of \( Q \), and \( \alpha = \frac{P_z}{Q_z} \) is a scalar factor which is the ratio of two depths of the corresponding feature points on two object feature planes.

The matrix \( \left[ R + \frac{(I - R)P_c + t}{d} n' \right] \) becomes the homography matrix in this projection model. In this paper it is denoted as \( H \) as following,

\[ H = \left[ R + \frac{(I - R)P_c + t}{d} n' \right]. \]  

(2-13)

As a result, Eq. (2-12) can be changed into

\[ \alpha q = Hp. \]  

(2-14)

This matrix encodes the mapping relationship among the three planes represented by \( \Pi_1, \Pi_2 \) and the Image Plane, under the moving object perspective projection model. Therefore, given each correspondence \( p \) and \( q \) between two images of the object at different positions, each pair of corresponding feature points results in two equations. Therefore, given four coplanar corresponding points between two positions, the normalized form of the homography matrix can be determined by solving the linear system.

In the following of this paper, the discussion will be mainly based on the second perspective projection model of moving object, as illustrated in Fig. 2.

Third, I will introduce how to form the normalized homography matrix by four
pairs of coplanar corresponding points. Let \( \mathbf{p}_i \) and \( \mathbf{q}_i \) be a pair of corresponding points on two images which are taken before and after the 3D motion, as shown in Fig. 2. Based on Eq. (2-14), there is:

\[
a_i \mathbf{q}_i = \mathbf{H} \mathbf{p}_i. \tag{2-14}
\]

The scalar factor \( a_i \) can be eliminated by describing the matrix \( \mathbf{H} \) in the following normalized form:

\[
\mathbf{H}_0 = \begin{pmatrix}
    h_1 & h_2 & h_3 \\
    h_4 & h_5 & h_6 \\
    h_7 & h_8 & 1
\end{pmatrix}.
\tag{2-15}
\]

\( \mathbf{H}_0 \) differs from \( \mathbf{H} \) by a uniform scalar factor; that is

\[
\mathbf{H} = \delta \cdot \mathbf{H}_0. \tag{2-16}
\]

Thus, by replacing \( \mathbf{H} \) with \( \mathbf{H}_0 \), Eq. (2-14) is still valid. By defining the homogeneous coordinate as

\[
\mathbf{P} = \begin{pmatrix}
    P_x \\
    P_y \\
    P_z
\end{pmatrix}, \quad \mathbf{Q} = \begin{pmatrix}
    Q_x \\
    Q_y \\
    Q_z
\end{pmatrix}, \quad \mathbf{p} = \omega \begin{pmatrix}
    p_x \\
    p_y \\
    1
\end{pmatrix}, \quad \mathbf{q} = \omega \begin{pmatrix}
    q_x \\
    q_y \\
    1
\end{pmatrix}
\tag{2-17}
\]

and eliminating the scalar factor \( a_i \), Eq. (2-14) becomes two equations:

\[
q_{xi} = \frac{p_{xi} h_1 + p_{yi} h_2 + h_3}{p_{xi} h_1 + p_{yi} h_8 + 1} \tag{2-18}
\]

\[
q_{yi} = \frac{p_{xi} h_4 + p_{yi} h_5 + h_6}{p_{xi} h_7 + p_{yi} h_8 + 1} \tag{2-19}
\]

Considering that \( n \) pairs of coplanar corresponding points on the two images which are taken before and after 3D motion are available to be used, the following linear system can be obtained:
Therefore, if at least four coplanar but non-collinear corresponding points are available to be used, the linear system will have a unique solution. If more than four coplanar corresponding points are given, a least mean squares technique can be used to solve Eq. (2-20) to obtain the normalized homography matrix $H_0$. Nowadays, many of libraries for various programming languages, such as C, C++, Java, and so on, can be used directly to obtain the normalized homography matrix $H_0$ from four pairs of coplanar corresponding points coordinate. For example, in OpenCV, the function “cvGetPerspectiveTransform” is one of convenient tools to obtain $H_0$.

2.2 Faugeras’ Homography Decomposition Method for 3D Motion Parameters Estimation

In Faugeras’ paper [32], he proposed a theory: “Eq. (2-8) has, in general, 8 different solutions. It has only 4 if and only if $H$ has a singular value of multiplicity 2. The problem is partially undetermined if and only if $H$ has a singular value of multiplicity 3.”

Using the singular value decomposition (SVD), $H$ can always be decomposed as:

$$H = UV^\dagger,$$

(2-21)

$\Sigma$ is a diagonal matrix, and $U$ and $V$ are orthogonal matrices (satisfying $U^U = V^V = I$). The elements of $\Sigma$ are the square roots of the eigenvalues of $HH^T$.

These eigenvalues $\lambda_i$ are positive and can be sorted in decreasing order: $\lambda_1 \geq \lambda_2 \geq \lambda_3$.

Using this decomposition, a new equation can be obtained:
\[ \Sigma = R_x + \frac{t_x n_i}{d_z} \]  

(2-22)

\( R, t \) and \( n \) are related to \( R_x, t_x \) and \( n_x \) by:

\[
\begin{align*}
R &= s U R_x V' \\
&= U t_x \\
n &= V n_x \\
d &= s \cdot d_z \\
s &= \det U \cdot \det V
\end{align*}
\]

(2-23)

Notice that \( R_x \) is a rotation matrix (i.e. \( R_x R_x^T = I \) and \( \det R_x = 1 \)).

Using the canonical basis \((e_1, e_2, e_3)\) and writing \( n_x = x_i e_i + x_2 e_2 + x_3 e_3 \), Eq. (2-22) gives us three vector equations:

\[ \lambda_i e_i = d_z R_x e_i + t_z \lambda_i (i = 1, 2, 3) \]  

(2-24)

Notice that, since \( n \) has a unit norm and \( V \) is orthogonal, \( n_x \) has also a unit norm \( \sum_{i=1}^{3} x_i^2 = 1 \). Eliminating \( t_z \) finally yields:

\[ d_z R_x (x_i e_i - x_j e_i) = \lambda_i x_i e_i - \lambda_j x_j e_j \quad (i \neq j) \]  

(2-25)

As \( R_x \) preserve the vector norm, the following set of equations can be obtained:

\[
\begin{align*}
\left(d_z^2 - \lambda_i^2\right) x_i^2 + \left(d_z^2 - \lambda_j^2\right) x_j^2 &= 0 \\
\left(d_z^2 - \lambda_i^2\right) x_i^2 + \left(d_z^2 - \lambda_j^2\right) x_j^2 &= 0 \\
\left(d_z^2 - \lambda_i^2\right) x_i^2 + \left(d_z^2 - \lambda_j^2\right) x_j^2 &= 0
\end{align*}
\]  

(2-26)

This can be considered as a linear system in the unknowns \( x_i^2, \ x_j^2 \) and \( x_j^2 \). As it must have a non-zero solution, its determinant must be zero:

\[ \left(d_z^2 - \lambda_i^2\right) \left(d_z^2 - \lambda_j^2\right) = 0 \]  

(2-27)

Therefore, the different cases can be obtained, according to the order of multiplicity of the singular values of \( H \):

1. \( \lambda_1 \neq \lambda_2 \neq \lambda_3 \) and \( d_z = \pm \lambda_2 \).
2. \( \lambda_1 = \lambda_2 \neq \lambda_3 \) or \( \lambda_1 \neq \lambda_2 = \lambda_3 \), and \( d_z = \pm \lambda_2 \).
3. \( \lambda_1 = \lambda_2 = \lambda_3 \) and \( d_z = \pm \lambda_2 \).
The solutions $d_{\Sigma} = \pm \lambda_1$ or $d_{\Sigma} = \pm \lambda_3$ are indeed impossible; let us prove it in case 1, for example. Assuming $d_{\Sigma} = \lambda_1$, the Eq. (2-26) yield: $x_1 = 0$ and $(\lambda_1^2 - \lambda_3^2)x_2^2 + (\lambda_1^2 - \lambda_2^2)x_3^2 = 0$. As $\lambda_1 \geq \lambda_2 \geq \lambda_3$, this implies $x_2 = x_3 = 0$, which is impossible because $n_{\Sigma}$ has a unit norm.

If $\lambda_1 \neq \lambda_3$, $x_1$, $x_2$ and $x_3$ can be expressed by using Eq. (2-26) and the constraint that $n_{\Sigma}$ is the vector with the unit norm:

$$\begin{cases}
x_1 = \epsilon_1 \sqrt{\frac{\lambda_1^2 - \lambda_3^2}{\lambda_2^2 - \lambda_3^2}} \\
x_2 = 0 \\
x_3 = \epsilon_3 \sqrt{\frac{\lambda_2^2 - \lambda_3^2}{\lambda_1^2 - \lambda_3^2}}
\end{cases} \quad (2-28)$$

The final number of solutions depends on the sign of $d_{\Sigma}$. However, there is no priori knowledge about the sign of $d_{\Sigma}$, in total eight solutions will be obtained when the singular values are distinct, four solutions will be obtained when two singular values are equal, and an indetermination will be obtained when the three are equal. But among these eight solutions to the problem, only two are possible by considering the physical interpretation of the results. The details will be explained in the next Section 2.5.

2.3 Zhang’s Homography Decomposition Method for 3D Motion Parameters Estimation

In this section, I will introduce Zhang’s [33] homography decomposition method for 3D motion parameters estimation, by which the closed-form solutions are obtained for recovering the rotation matrix $R$, normalized translation vector $t_0$ and the normal vector $n$ of the planar object on the position before the 3D motion from $H_0$.

I will start with $H_0$ as defined in Eq. (2-15), because $H_0$ can be obtained easily by many methods. $H_0$ and $H$ (as defined in Eq. (2-13)) are related by an unknown scalar factor $\phi$. 

20
Define the following:
\[ t_0 = R' \left( \frac{(I - R)P_c + t}{d} \right) \]  
\[ k^2 = t_0 \cdot t_0, \quad k > 0 \]  
\[ m = n \cdot t_0 \] 

Now there is:
\[ \delta \cdot H_0 = \begin{bmatrix} R + \frac{(I - R)P_c + t}{d} n' \end{bmatrix} \]  

And by multiplying the transpose of Eq. (2-33) and itself, Eq. (2-34) will be obtained:
\[ \delta^2 H_0' H_0 = I + nt_0' + t_0 n' + k^2 nn' \]  

Now it is necessary to find the analytical form of the three eigenvalues and their corresponding eigenvectors. For the general situation, it is possible to assume that \( t_0 \) and \( n \) are not in the same direction. By the definitions of the eigenvalues and eigenvectors, It is clear that \( \lambda_2 = 1 \) is one of solutions of the eigen equation Eq. (2-35):
\[ \begin{cases} \delta^2 H_0' H_0 - \lambda I = 0 \\ (1 - \lambda)I + nt_0' + t_0 n' + k^2 nn' = 0 \end{cases} \]  

It is easy to determine the eigenvector for \( \lambda_2 = 1 \) as:
\[ v_2 = t_0 \times n \]  

Since \( \delta^2 H_0' H_0 \) is symmetric, as indicated in Eq. (2-34), the other two eigenvectors must be on the plane perpendicular to \( v_2 \). It follows that the other two eigenvectors can be assumed to be of the form:
\[ v_{1,3} = at_0 + bn \]  

Thus, from the relation of the eigenvalue and eigenvector, the next equation must be satisfied:
\[ \delta^2 H_0' H_0 (at_0 + bn) = (a + am + b)t_0 + (amk^2 + ak^2 + bk^2 + bm + b)n \]  
\[ = \lambda_{1,3} (at_0 + bn) \]
Here, \( \lambda_{1,3} \) are the two corresponding eigenvalues of \( v_{1,3} \). From Eq. (2-38), two equations can be obtained:

\[
\begin{align*}
\hat{\lambda}_{1,3}a &= a + am + b \\
\hat{\lambda}_{1,3}b &= amk^2 + ak^2 + bk^2 + bm + b \\
\end{align*}
\] (2-39)

With the assumption that \( \hat{\lambda}_{1,3}, a \) and \( b \) are unknown, \( k \) and \( m \) are known, the solutions of Eq. (2-39) can be obtained. In summary, the general situation is described as follows \((m \neq -1)\):

\[
\begin{align*}
\hat{\lambda}_1 &= 1 + m + \frac{2k(m+1)}{-k + \sqrt{k^2 + 4(m+1)}} \\
\hat{\lambda}_3 &= 1 + m + \frac{2k(m+1)}{-k - \sqrt{k^2 + 4(m+1)}} \\
\end{align*}
\] (2-40)

and

\[
\begin{align*}
v_1 &= \frac{-k + \sqrt{k^2 + 4(m+1)}}{2k(m+1)} t_0 + n \\
v_3 &= \frac{-k - \sqrt{k^2 + 4(m+1)}}{2k(m+1)} t_0 + n \\
\end{align*}
\] (2-41)

If \( m = -1 \),

\[
\begin{align*}
\hat{\lambda}_1 &= k^2 \\
\hat{\lambda}_3 &= 0 \\
\end{align*}
\] (2-42)

and

\[
\begin{align*}
v_1 &= \frac{1}{k^2} t_0 + n \\
v_3 &= t_0 \\
\end{align*}
\] (2-43)

It is easy to verify that \( v_1 \cdot v_3 = 0 \).

To simplify the mathematical expressions, the following parameters will be defined:

\[
\begin{align*}
\varphi &\overset{\text{def}}{=} \frac{-k + \sqrt{k^2 + 4(m+1)}}{2k(m+1)} \\
\theta &\overset{\text{def}}{=} \frac{-k - \sqrt{k^2 + 4(m+1)}}{2k(m+1)} \\
\end{align*}
\] (2-44)

Now the eigenvalues and eigenvectors in case \( m \neq -1 \) can be rewritten as following:
\[
\begin{align*}
\mathbf{v}_1 &= \varphi \cdot \mathbf{t}_0 + \mathbf{n} \\
\mathbf{v}_3 &= \theta \cdot \mathbf{t}_0 + \mathbf{n}
\end{align*}
\]  

(2-45)

Based on the solutions of the obtained eigenvalues, it is possible to examine the various cases for the three eigenvalues. By the definitions of \( k \) and \( m \) in Eq. (2-31) and Eq. (2-32), it is obvious that the following two constraints are always valid:

\[
k^2 \geq p
\]  

(2-47)

\[
\|k\| \geq \|p\|
\]  

(2-48)

Because \( \delta^2 \mathbf{H}_0^\dagger \mathbf{H}_0 \) is a real symmetric matrix, it always has real eigenvalues. Therefore, a third constraint is:

\[
k^2 + 4(p + 1) \geq 0
\]  

(2-49)

which means that \( p \geq -\frac{k^2}{4} - 1 \).

Based on the above three constraints, it can be shown that

\[
\lambda_1 \geq \lambda_2 = 1
\]  

(2-50)

where equality holds if and only if \( p = -2, k = 2 \) or \( p = -1, k = 1 \); and

\[
\lambda_3 \leq \lambda_2 = 1
\]  

(2-51)

and equality holds if and only if \( p = k \).

Note that those conditions that make the above constraints become equalities violate the assumption that \( \mathbf{t}_0 \) and \( \mathbf{n} \) are not aligned. Therefore, in the general case that the two vectors are not aligned, a strict inequality can be obtained:

\[
\lambda_1 > \lambda_2 = 1 > \lambda_3
\]  

(2-52)

The following discussion gives an analytical closed-form solution for the decomposition of the \( \mathbf{H}_0 \) matrix in the general case.

Until now, the analytical form of the three eigenvalues and eigenvectors of the
matrix $H'HH$ have been obtained. In another way, it is possible to obtain their numerical forms by the using SVD of the $H_0$. Assuming that three singular values of $H_0$ are $\rho_1 > \rho_2 > \rho_3$, then their corresponding eigenvectors are $\pm v_1^*, \pm v_2^*, \pm v_3^*$. Since
\[
\delta \rho_2' = \lambda_2 = 1
\] (2-53)
the scalar factor can be determined as
\[
\delta = \frac{1}{\rho_2'}
\] (2-54)
and the other two singular values of $H$, $\rho_1$ and $\rho_3$, can be obtained as
\[
\rho_1 = \delta \rho_1'
\] (2-55)
\[
\rho_3 = \delta \rho_3'
\] (2-56)
By using the relationship between the singular values and the eigenvalues of the same matrix ($\lambda = \rho^2$), two equations from Eq. (2-40) can be obtained as following:
\[
\begin{cases}
\lambda_1 = \rho_1^2 \\
\lambda_3 = \rho_3^2
\end{cases}
\] (2-57)
Solving these two equations, there are:
\[
\begin{cases}
k = \rho_1 - \rho_3 \\
m = \rho_1 \rho_3 - 1
\end{cases}
\] (2-58)
Now by defining:
\[
\mu = \|v_1\| = \sqrt{(\varphi \cdot t_0 + n)^2 (\varphi \cdot t_0 + n)} = \sqrt{\varphi^2 k^2 + 2 \varphi m + 1}
\] (2-59)
and
\[
\sigma = \|v_3\| = \sqrt{(\theta \cdot t_0 + n)^2 (\theta \cdot t_0 + n)} = \sqrt{\theta^2 k^2 + 2 \theta m + 1}
\] (2-60)
Thus,
\[
\pm v_1^* = \frac{v_1}{\|v_1\|} = \frac{\varphi \cdot t_0 + n}{\mu}
\] (2-61)
and

$$\pm v_3^* = \frac{v_3}{\|v_3\|} = \frac{\theta \cdot t_0 + n}{\sigma}$$  \hspace{1cm} (2-62)

By solving the four sets of the equations above, finally the next four solutions can be acquired:

$$\begin{cases}
    t_0 = \frac{\mu \cdot v_1 - \sigma \cdot v_3}{\varphi - \theta} \\
    n = \frac{\varphi \sigma \cdot v_3 - \theta \mu \cdot v_1}{\varphi - \theta}
\end{cases}$$  \hspace{1cm} (2-63)

$$\begin{cases}
    t_0 = \frac{\mu \cdot v_1 - \sigma \cdot v_3}{\varphi - \theta} \\
    n = \frac{\varphi \sigma \cdot v_3 - \theta \mu \cdot v_1}{\varphi - \theta}
\end{cases}$$  \hspace{1cm} (2-64)

$$\begin{cases}
    t_0 = \frac{\mu \cdot v_1 + \sigma \cdot v_3}{\varphi - \theta} \\
    n = \frac{\varphi \sigma \cdot v_3 + \theta \mu \cdot v_1}{\varphi - \theta}
\end{cases}$$  \hspace{1cm} (2-65)

$$\begin{cases}
    t_0 = \frac{\mu \cdot v_1 + \sigma \cdot v_3}{\varphi - \theta} \\
    n = \frac{\varphi \sigma \cdot v_3 + \theta \mu \cdot v_1}{\varphi - \theta}
\end{cases}$$  \hspace{1cm} (2-66)

Eq. (2-63), Eq. (2-64), Eq. (2-65) and Eq. (2-66) are the decomposition solutions of the homography matrix $H_0$. It is easily to obtain the rotation matrix $R$ by estimating the results of the translation vector $t_0$ and the plane normal vector $n$ in the traditional way:

$$R = \delta H_0 (I + t_0 n')^{-1}$$  \hspace{1cm} (2-67)

As there are four sets of $t_0$ and $n$ respectively, four estimated $R$s will be realized finally.

Because Zhang’s method is more advanced than Faugeras and Lustman’s, and the computer cost is lower than the latter, my research mainly focuses on Zhang’s
method. However, the estimated translation vector $t_0$ is just the normalized translation vector which is defined by Eq. (2-30). In order to obtain the real translation vector $t$ in the moving object perspective projection model, the coordinate of rotation center $P_c$ which is the geometry center on the object before the 3D motion is necessary. The common method proposed in the Zhang’s paper [33], are suitable to the situation which is under the two cameras perspective projection model. Therefore, another total new method will be proposed to estimate the coordinate of rotation center $P_c$ in the next section.

### 2.4 Iterative Method to Estimate the Real Rotation Center Coordinate on the Object

![Image](image.png)

Fig. 3 Geometry relationship of object plane ($\Pi_1$) and virtual plane ($\Omega_1$)
Under the perspective projection model of the moving object, the coordinate of the rotation center (the average coordinate of all feature points in 3D space) is a key factor to estimate the real translation vector $t$ from the normalized translation vector $t_0$. As a result, the exact coordinates on the image of the rotation center projected from the front pose on different directions are necessary to improve the estimation accuracy.

In this section, a kind of iterative method will be introduced, which are used in my research, to estimate the real rotation center image coordinate.

First, I assume that there is a virtual plane $\Omega_1$, which has the same rotation center with the object plane $\Pi_1$, which is before 3D motion. Meanwhile, $\Omega_1$ is perpendicular to the optical axis from the camera.

Fig. 3 illustrates the geometry relationship between $\Pi_1$ and $\Omega_1$. $\{A_1, B_1\}$ is feature points of the virtual plane $\Omega_1$. X axis is parallel to the image plane. Those projected feature points on the image plane, $\{a_1, b_1\}$, are the projected points of the feature points on both object plane and virtual plane respectively. Z’ axis is the direction from the origin (camera) pointing to the rotation center in the 3D space. I make it the optical axis.
With the feature points coordinated on the image, it is necessary to estimate the image coordinate of the rotation center $o_1$, because the intersection point between the line pointing to the object center and the image plane does not coincide with the average of projected feature points.

As the illustration shown in Fig. 3, it is obvious that there is only one plane ($A1B1$) which is perpendicular to $oo_1$. Meanwhile, this plane intersects the image plane on the line which passes through the rotation center. Therefore, the procedure to get the projected rotation center is like bellow:

**[Estimation algorithm for the rotation center on the image plane]**

1. The initial value $o_1$ of the projected point of the rotation center is assumed as the average coordinate of the projected feature points $\{a_1, b_1\}$ on the image plane.
(2) The direction vector $\overrightarrow{o o_j}$ from the camera $o$ to the current projected point $o_j$ of the rotation center can be known. Here, $j (=1, 2, \ldots)$ is the index of iteration.

(3) The plane $(a_j' b_j')$ which is perpendicular to the direction vector $\overrightarrow{o o_j}$ and passes through $o_j$ is created by Eq.(2-68).

$$A \cdot X + B \cdot Y + C \cdot Z - (A \cdot X_c + B \cdot Y_c + C \cdot Z_c) = 0 \quad (2-68)$$

Here, $(X_c, Y_c, Z_c)'$ is the coordinate of the current projected point of the rotation center; $(A, B, C)'$ is the direction vector $\overrightarrow{o o_j}$ from the camera $o$ to the current projected point of the rotation center $o_j$.

(4) By using the vectors from the camera $o$ to the projected the feature points $\{a_1', b_1\}$, the intersection points $\{a_j', b_j'\}$ between these vectors and the plane $(a_j' b_j')$ obtained in step 3 can be calculated.

(5) The average point $o_j'$ of the intersection points $\{a_j', b_j'\}$ can be calculated. Note that the average point is not on the image plane.

(6) If the center point $o_j'$ of the intersection points is very close to the current projected point $o_j$ of the rotation center, the iteration finishes. The stop condition in this iteration is when the distance between these two points is less than $10^{-6}$ pixel.

Otherwise, a new direction vector $\overrightarrow{o o_j'}$ from the camera $o$ to this center point $o_j'$ is calculated, and it is possible to obtain the intersection point $o_{j+1}$ between direction vector $\overrightarrow{o o_j'}$ and image plane. Make $o_{j+1}$ as a new projected point of the rotation center, and go back to step 2.

By the algorithm above, the image coordinate of rotation center $o_2$ can be obtained. Although, the proposed method is a kind of iterative method, the experiment proves that this algorithm can converge to the real solution very fast. Therefore, it is possible to estimate the rotation center coordinate on the image accurately and quickly. The detail about the experiments result will be described in Chapter 5.
By estimating the coordinate of rotation center on the image, it is possible to obtain the 3D space coordinate $\frac{P}{d}$ of the rotation center without the depth. Then by the inverse operation of Eq. (2-30), the real translation vector $\frac{t}{d}$ without the depth as:

$$\frac{t}{d} = R_{t_0} - (I - R) \frac{P}{d}$$

(2-69)

Therefore, under the moving object perspective projection model, the analytical closed-form solution of the 3D motion parameters (3D rotation matrix $R$, real 3D translation vector $t$ and the normal vector $n$ of the object plane on the position before 3D motion) can be obtained by homography decomposition method proposed by Zhang in 1990s [33]. The computation cost is extremely controlled than the non-linear methods, such as Gauss-Newton method. Comparing to the homography decomposition method proposed by Faugeras, Zhang’s method is more flexible. It is very simple to modify the algorithm to fit the different perspective projection models. Although various situations were discussed in Zhang’s paper to make the method more accurate and more comprehensive, my research will mainly be focused on the common situation which is discussed above.

### 2.5 Solution Ambiguity Problem of 3D Motion Parameters Estimation by Homography Decomposition

By the methods which are introduced in Section 2.2 and Section 2.3, numerical solutions will be obtained in estimating 3D motion parameters by homography decomposition. By Faugeras’ method, 8 sets of solutions will be obtained, Eq. (2-28), and By Zhang’s method, 4 sets of solutions are inevitable, Eqs. (2-63) to (2-66).
Because every set of estimated solution must contain the three 3D motion parameters: $\mathbf{R}$, $\mathbf{t}_0$ and $\mathbf{n}$. Therefore, when “the solution” is mentioned in this paper, it only means the normal vector $\mathbf{n}$ of $\Pi_1$.

Considering Zhang’s method, the four estimated normal vectors of object plane $\Pi_1$ before the 3D motion will be displaced as illustrated in Fig. 5. Fig. 5 illustrates one of possible situations of these four estimated normal vectors. As shown in Fig. 5, $\mathbf{n}_1$ and $\mathbf{n}_2$ are on a straight line and in the opposite directions. So are $\mathbf{n}_3$ and $\mathbf{n}_4$.

Recalling that the solution should be “real” in the sense that all feature points $\{\mathbf{P}_i\}$ should satisfy the general constraint of $\mathbf{n}'\mathbf{P}_i > 0$ (which means the feature plane should be in front of the camera), two of the four sets of solutions in Eq.s (2-63) to (2-66) in Section 2.3 can be easily removed, i.e. only those two solutions for which
One question that remains is whether or not these two solutions can be distinguished so that the “true” solution can be uniquely identified. With any a priori knowledge, the only recourse is to check the distribution of the two solutions in terms of the chirality constraint [39] with respect to the reference plane, i.e. check whether or not the two solutions are on the same side of the plane.

In the absence of a priori information to guide the choice, Faugeras proposed three kinds of ways to proceed [32]:

- Look at a second plane (add a reference plane).
- Use a third image (increase a view point).
- Use known geometric relationships between tokens in the plane, for example that two lines are orthogonal (increase the necessary information).

In the first two cases, two pairs \((n_a, n_b)\) and \((n'_a, n'_b)\) of solutions are obtained, and then it is necessary to find a compatible pair \((n_j, n'_j)\), i.e. find a common plane equation by looking at a single plane from three points, or find a common motion by looking at two planes from two positions. In general, there is only one compatible pair, and the problem therefore has a unique solution.

In the third case, it is possible to set two pairs of corresponding lines between two images, given by their equations. Let \(R_1\) and \(R_2\) the two estimated rotation matrix for the solution. The direction of the reconstructed lines for a rotation \(R\) can be cleared. If we know the angle between these two vectors, checking this angle for \(R_1\) and \(R_2\) will, in general, give the right solution.

In a word, Faugeras intended to increase the information to determine the unique real solution.

Another way, Zhang proposed his theoretical argument on the ambiguous solution problem [38].
The necessary and sufficient condition for the case that the two solutions lie on opposite sides of the reference plane [39] (i.e. the camera center goes to the other side of the reference plane after 3D motion) is:

\[ t_0 \cdot n < -1 \]  \hspace{1cm} (2-70)

Given Eqs. (2-63) to (2-66), for both solutions \( i = 1,2 \), it can be shown that:

\[ t_{0i} \cdot n_i = m \]  \hspace{1cm} (2-71)

which again is consistent with the definition in Eq. (2-32). This result means that if \( m < -1 \), the solution straddle the reference plane; otherwise, both solution are on the same side of the reference plane. Consequently, the two solutions are intrinsically indistinguishable. Therefore, this ambiguity of the two solutions cannot be resolved if there is no further a priori information available.

It is worth nothing that if the condition \( t_{0i} \cdot n_i < -1 \) is valid, then this is the case that for both solutions, the object crosses the reference plane after 3D motion. This case is possible if and only if the reference plane is “transparent”, i.e. all the features on this plane can be viewed from both sides of it. In this case the homography mapping is still valid, and thus, the recovered solutions are also valid.

In a word, Zhang had proved that the ambiguous estimated solutions are inevitable by his theory.

In summary, the other two cannot be distinguished without any other a priori constraints. Therefore, when we estimate the 3D motion parameters from only two pieces of images, we will always obtain ambiguous solutions by the method of homography decomposition. This is the famous “Solution Ambiguity Problem”.

The inevitability of the ambiguity problem to estimate solutions has been widely accepted for nearly 30 years. It is considered to be one of the basic properties and mathematical defects in homography decomposition, even today [12, 13, 15, 37]. Many researchers are focus on the researched to propose the suitable constraint conditions by which the ambiguous solution can be uniquely obtained.
However, my research is focus on the intrinsic reason of the solution ambiguity problem. I found that the ambiguity problem to estimate solutions is not inevitable; even only two pieces of images are available and no more extra information to estimate the 3D motion parameters by homography matrix. The details will be discussed in the next chapters.
Even though the solution ambiguity problem seems inevitable, some researchers have suggested the possibility of a unique solution and a kind of theoretical proof was demonstrated by Longuet-Higgins in the 1980s [40].

Longuet-Higgins derived the proof of the unique solution under two important assumptions: \( t_0' t_0 = 1 \), \( n' P = 1 \). Meanwhile, by rotating some vectors with eigenvector matrix \( V \) of \( HH' \): \( n' = V' n \), \( t_0' = V' t_0 \), \( P' = V' P \), \( p' = V' p \), he obtained:

\[
t_0' = (t_1', \quad 0, \quad t_3') \quad n' = (n_1', \quad 0, \quad n_3')
\]  \( (3-1) \)

Because the real solution should always be in front of the camera, the feature points on the images should also be in front of the camera, \( n' p > 0 \). Taking into consideration the rotation by \( V \) and Eq. (3-1), the general reality constraint can be expressed as

\[
n_1' p_x' + n_3' p_z' > 0
\]  \( (3-2) \)

Because the sign (+ or -) of \( n_1' \) and \( n_3' \) cannot be determined, there are four kinds of possible combinations. However, if \( n_1' p_x' > n_3' p_z' \) for all the feature points, reversing the sign of \( n_3' p_z' \) will not violate the reality constraint in Eq. (3-2); however, reversing the sign of \( n_1' p_x' \) will. Alternatively, if \( n_1' p_x' < n_3' p_z' \), reversing the sign of \( n_1' p_x' \) will not violate the reality constraint in Eq. (3-2); however, reversing the sign of \( n_3' p_z' \) will. In either case, an ambiguous solution will occur; otherwise a unique solution will be obtained.

Therefore, the solution ambiguity problem will occur, if and only if \( n_1' p_x' - n_3' p_z' \) has the same sign for all the visible feature points. Otherwise, it is possible that a unique solution will be obtained. (For details, refer to Longuet-Higgins
As the conclusion, Longuet-Higgins claimed:

“The uniqueness of the interpretation, for the case in which some of the visible points are nearer to \( O \) and others nearer to \( O' \), corresponds to the uniqueness of the interpretation of the optic flow field of a planar region part of which lies ‘ahead’ and part ‘astern’ relative to the observer’s motion.”

Although Longuet-Higgins proposed one kind of possibility for a unique solution obtained by homography decomposition, this property of homography decomposition was not accurately described. The proof was inadequate because the whole procedure in proving it involved two assumptions. The inadequacy of Longuet-Higgins’ proof led it lack of reliability.

Another way, Faugeras [32] had also mentioned the unique solution in estimating the 3D motion parameters by the homography decomposition. It was very similar to what Longuet-Higgins proposed before, but more clearly described under the two cameras perspective projection model. However, he just described that was the interesting property. Not any further description or the theoretical proof was available in his paper.

During more than 30 years until now, there are quite rare researches on this topic. Longuet-Higgins [40] and Faugeras [32] are the only two we can find nowadays who had noticed the avoidability of the solution ambiguity problem in estimating 3D motion parameters by homography decomposition. And either of these two research were not able to proposed the sufficiently theoretical proof and the completely description. As a result, the solution ambiguity problem has been considered the basic property of the homography decomposition in estimating the 3D motion parameters.

In my research, I focus on this traditional misunderstanding. Meanwhile, I will not only propose more complete and accurate descriptions of the properties of the solution ambiguity problem in homography decomposition, but also verify each theory I propose by using a different method from that by Loguet-Higgins, which is under more common constraints. As a result, I intent to create a more theoretical and more
complete description on the property of homography decomposition method in estimating 3D motion parameters.

First of all, in my research, I set the condition to a very extreme situation as the premise, which the useful information to estimate the 3D motion parameters, or to obtain the homography matrix is quite limited. Under this condition, the solution ambiguity problem will absolutely inevitable by the general knowledge on the homography decomposition method in estimating the 3D motion parameters.

**[Premise: Condition setting in my research]**

Only two pieces of images are available to estimate the 3D motion parameters. There is absolutely not any other extra information or constraints can be used.

Even under this condition, solution ambiguity problem on homography decomposition shows two kinds of dependences. These dependences become the topic of this chapter.

### 3.1 Dependence on 3D Motion Parameters of Solution Ambiguity by Homography Decomposition

This section describes one of the properties of solution ambiguity obtained with homography decomposition in estimating the 3D motion parameters when the object feature region size is fixed. Meanwhile, I will geometrically and theoretically prove what I propose in this section.

#### 3.1.1 Dependence on 3D Motion Parameters and its Theoretical Proof

In this section, I will not only propose a kind of dependence of the solution ambiguity in estimating the 3D motion parameters by homography decomposition when the object feature region size is fixed, but also presents a new strictly theoretical proof, which is under a more common constraint than that in Longuet-Higgins’ work [40].
By substituting the first and the third singular values \((\rho_1, \rho_3)\) into the expressions of estimated normal vector \(\mathbf{n}\) of the initial feature plane \(\Pi_1\), i.e. object plane which is in the position before the 3D motion, as in the Eqs. (2-63) to (2-66), four normal vector \(\mathbf{n}\) expressions with \(\rho_1\) and \(\rho_3\) will be obtained as following:

\[
\mathbf{n} = \begin{cases} 
\frac{1 - \rho_3^2}{\rho_1^2 - \rho_3^2} \mathbf{v}_3 + \frac{\rho_1^2 - 1}{\rho_1^2 - \rho_3^2} \mathbf{v}_1 \quad \mathbf{n}_1 \\
\frac{1 - \rho_3^2}{\rho_1^2 - \rho_3^2} \mathbf{v}_3 + \frac{\rho_1^2 - 1}{\rho_1^2 - \rho_3^2} \mathbf{v}_1 \quad \mathbf{n}_2 \\
\frac{1 - \rho_3^2}{\rho_1^2 - \rho_3^2} \mathbf{v}_3 - \frac{\rho_1^2 - 1}{\rho_1^2 - \rho_3^2} \mathbf{v}_1 \quad \mathbf{n}_3 \\
\frac{1 - \rho_3^2}{\rho_1^2 - \rho_3^2} \mathbf{v}_3 - \frac{\rho_1^2 - 1}{\rho_1^2 - \rho_3^2} \mathbf{v}_1 \quad \mathbf{n}_4 
\end{cases}
\]  

Considering the relation between eigenvalues and singular values \((\lambda_i = \rho_i^2)\), it is possible to substitute \(\lambda_1\) and \(\lambda_3\) into Eq. (3-3), and it can be rewritten into

\[
\mathbf{n} = \begin{cases} 
\frac{1 - \lambda_3}{\lambda_1 - \lambda_3} \mathbf{v}_3 + \frac{\lambda_1 - 1}{\lambda_1 - \lambda_3} \mathbf{v}_1 \quad \mathbf{n}_1 \\
\frac{1 - \lambda_3}{\lambda_1 - \lambda_3} \mathbf{v}_3 + \frac{\lambda_1 - 1}{\lambda_1 - \lambda_3} \mathbf{v}_1 \quad \mathbf{n}_2 \\
\frac{1 - \lambda_3}{\lambda_1 - \lambda_3} \mathbf{v}_3 - \frac{\lambda_1 - 1}{\lambda_1 - \lambda_3} \mathbf{v}_1 \quad \mathbf{n}_3 \\
\frac{1 - \lambda_3}{\lambda_1 - \lambda_3} \mathbf{v}_3 - \frac{\lambda_1 - 1}{\lambda_1 - \lambda_3} \mathbf{v}_1 \quad \mathbf{n}_4 
\end{cases}
\]  

Even though only two pieces of image are available to estimate the 3D motion parameters, there still one fact can be used as the constraint to eliminate the false decomposition solutions. The only constraint which can be used to distinguish the real solution from the four candidates in Eq. (3-4) obtained with Zhang and Hanson's method [33] is the fact:

The solution should be “real” in the sense. It means that all of the feature points on the object plane should be in front of the camera. It is possible to describe this fact by the algebraic expression as: to all feature points \(\{\mathbf{P}_i\}\), the real estimated normal vector \(\mathbf{n}\) should satisfy the general constraint
\[ n^T P_i > 0 \]  

This fact is the constraint of “solution reality”, and it is the only constraint can be used under the extreme situation to estimate the 3D motion parameters.

Therefore, the inner product between \( n_j \) (\( j = 1, 2, 3, 4 \)) and \( P_i \) (\( i \) is the index of the feature points) needs to be calculated. As a result, \( n_j^T P_i \) can be expressed as

\[
\begin{align*}
    n_1^T P_i &= \frac{1 - \lambda_3}{\lambda_1 - \lambda_3} v_3^T P_i + \frac{\lambda_1 - 1}{\lambda_1 - \lambda_3} v_i^T P_i \\
    n_2^T P_i &= -\left( \frac{1 - \lambda_3}{\lambda_1 - \lambda_3} v_3^T P_i + \frac{\lambda_1 - 1}{\lambda_1 - \lambda_3} v_i^T P_i \right) \\
    n_3^T P_i &= -\left( \frac{1 - \lambda_3}{\lambda_1 - \lambda_3} v_3^T P_i - \frac{\lambda_1 - 1}{\lambda_1 - \lambda_3} v_i^T P_i \right) \\
    n_4^T P_i &= \frac{1 - \lambda_3}{\lambda_1 - \lambda_3} v_3^T P_i - \frac{\lambda_1 - 1}{\lambda_1 - \lambda_3} v_i^T P_i
\end{align*}
\]

The number of the normal vector \( n \) which can satisfies the constraint \( P_i \cdot n > 0 \) for all the feature points, will be the number of the solution to the homography decomposition method. Thus, if the linear estimation has a unique solution, there will be only one candidate \( n \) that satisfies \( P_i \cdot n > 0 \) for all the feature points.

To determine whether the inner product between \( P_i \) and \( n_j \) is positive or negative, it is necessary to consider the next conditions from the real 3D motion model. From Fig. 2, it is obvious that there are two kinds of situations for the distances between camera \( O \) and the corresponding feature points before and after 3D motion, i.e. the feature point becomes near to the camera or far from the camera:

\[
\begin{align*}
    &P'P > Q'Q \\
    &P'P < Q'Q
\end{align*}
\]

The situation that the distance between camera \( O \) and the corresponding feature point does not changed during the 3D motion, i.e. \( P'P = Q'Q \), will not be discussed here, because it represents extremely special circumstances.

The homography mapping between \( \Pi_1 \) and \( \Pi_2 \) can be represented as

\[ Q_j = HP_i \]  

39
Also, $H$ can be decomposed by SVD,

$$H = U\Sigma V.'$$

(3-9)

Combine Eq. (3-8) and Eq. (3-9), as a result, the distance between $Q_i$ and $O$, $Q_i'O_i$, can be expressed as

$$Q_i'O_i = P_i'H'HP_i$$

$$= P'_i(\Sigma^2 V'P_i$$

$$= P'_iV_i \Sigma^2 V'P_i$$

$$= P'_i(v_1, v_2, v_3) \begin{pmatrix} \rho_1^2 & \rho_2^2 & \rho_3^2 \\ v_1^T & v_2^T & v_3^T \end{pmatrix} P_i$$

(3-10)

Considering that for the middle eigenvalue of matrix $H'H$ is equal to 1, i.e. $\lambda_2 = 1$, Eq. (3-10) can be rewritten into

$$Q_i'O_i = \lambda_i(v_1^T P_i)^2 + (v_2^T P_i)^2 + \lambda_3(v_3^T P_i)^2$$

(3-11)

Here, $v_1$, $v_2$ and $v_3$ are the component vectors of right singular vector matrix $V$, and $\lambda_i$ is the eigenvalue of $H'H$.

Another way, because $V$ is the orthogonal matrix, $\|P_i\|$, i.e. $P_i^TP_i$, does not change, even if $P_i$ is rotated by $V$,

$$P_i^TP_i = (V^TP_i)^T(V^TP_i)$$

$$= (v_1^T P_i)^2 + (v_2^T P_i)^2 + (v_3^T P_i)^2$$

(3-12)

Recalling the two kinds of situations for the distances between camera $O$ and the corresponding feature points before and after 3D motion, which has been discussed above, by substituting Eq. (3-11) and Eq. (3-12) into Eq. (3-7), two more constraints can be obtained:
\[ P'P > Q'Q \]
\[(v'_1P)^2 + (v'_2P)^2 > \lambda_1(v'_1P)^2 + \lambda_3(v'_3P)^2, \quad (3-13)\]
\[(1 - \lambda_3)(v'_3P)^2 > (\lambda_1 - 1)(v'_1P)^2 \]

or

\[ P'P < Q'Q \]
\[(v'_1P)^2 + (v'_2P)^2 < \lambda_1(v'_1P)^2 + \lambda_3(v'_3P)^2 \]
\[(1 - \lambda_3)(v'_3P)^2 < (\lambda_1 - 1)(v'_1P)^2 \]

From Eq. (3-13) and Eq. (3-14), it will be possible to clarify whether the inner product between \( P_i \) and \( n_j \) is positive or negative, by substituting \( v_3^iP_i \) and \( v_1^iP_i \) into any root sign of Eq. (3-6).

However, without any other constraints or the prior knowledge, the sign (+ or -) of \( v_3^iP_i \) and \( v_1^iP_i \) can be arbitrarily combined, i.e. \( v_3^iP_i \) and \( v_1^iP_i \) can both be positive, or they also can both be negative, or \( v_3^iP_i \) is positive and \( v_1^iP_i \) is negative, or otherwise. There are total four kinds of combinations of \( v_3^iP_i \) and \( v_1^iP_i \) signs. Therefore, to each feature point \( P \), there will be four possible combinations of sign for \( v_1^iP_i, v_3^iP_i \), as summarized in Table 1.

<table>
<thead>
<tr>
<th>Case</th>
<th>( v_1^iP_i )</th>
<th>( v_3^iP_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>II</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>III</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>IV</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

By using Eq. (3-13) or Eq. (3-14), it is possible to determine the sign of \( n_j^iP_i \) for each possible case ( I , II , III, and IV). As a result, there are totally eight kinds of situations. All of the possible situations will be discussed case by case following.
First, when \( P'P > Q'Q \), Eq. (3-13) will be the constraint:

**Constraint 1:** \((1 - \lambda_3)(v'_1 P)^2 > (\lambda_1 - 1)(v'_i P)^2\)

**Case I:**

\[
\begin{align*}
\begin{cases}
 v'_1 P > 0 \\
 v'_3 P > 0
\end{cases}
\end{align*}
\]

\[
\begin{align*}
 n'_1 P &= \sqrt{\frac{(1 - \lambda_3)(v'_1 P)^2}{\lambda_1 - \lambda_3} + \frac{(\lambda_1 - 1)(v'_i P)^2}{\lambda_1 - \lambda_3}} > 0 \\
n'_2 P &= -\sqrt{\frac{(1 - \lambda_3)(v'_1 P)^2}{\lambda_1 - \lambda_3} - \frac{(\lambda_1 - 1)(v'_i P)^2}{\lambda_1 - \lambda_3}} < 0 \\
n'_3 P &= -\sqrt{\frac{(1 - \lambda_3)(v'_1 P)^2}{\lambda_1 - \lambda_3} + \frac{(\lambda_1 - 1)(v'_i P)^2}{\lambda_1 - \lambda_3}} < 0 \\
n'_4 P &= \sqrt{\frac{(1 - \lambda_3)(v'_1 P)^2}{\lambda_1 - \lambda_3} - \frac{(\lambda_1 - 1)(v'_i P)^2}{\lambda_1 - \lambda_3}} > 0
\end{align*}
\]

Therefore, the four signs of \( n'_i P \) in this situation will be \((+, -, -, +)\). And two normal vectors are the candidates of the final estimation solution: \( n_1 \) and \( n_4 \).

**Case II:**

\[
\begin{align*}
\begin{cases}
 v'_1 P > 0 \\
 v'_3 P < 0
\end{cases}
\end{align*}
\]

\[
\begin{align*}
 n'_1 P &= -\sqrt{\frac{(1 - \lambda_3)(v'_1 P)^2}{\lambda_1 - \lambda_3} + \frac{(\lambda_1 - 1)(v'_i P)^2}{\lambda_1 - \lambda_3}} < 0 \\
n'_2 P &= \sqrt{\frac{(1 - \lambda_3)(v'_1 P)^2}{\lambda_1 - \lambda_3} - \frac{(\lambda_1 - 1)(v'_i P)^2}{\lambda_1 - \lambda_3}} > 0 \\
n'_3 P &= \sqrt{\frac{(1 - \lambda_3)(v'_1 P)^2}{\lambda_1 - \lambda_3} + \frac{(\lambda_1 - 1)(v'_i P)^2}{\lambda_1 - \lambda_3}} > 0 \\
n'_4 P &= -\sqrt{\frac{(1 - \lambda_3)(v'_1 P)^2}{\lambda_1 - \lambda_3} - \frac{(\lambda_1 - 1)(v'_i P)^2}{\lambda_1 - \lambda_3}} < 0
\end{align*}
\]

Therefore, the four signs of \( n'_i P \) in this situation will be \((--, +, +, -)\). And two normal vectors are the candidates of the final estimation solution: \( n_2 \) and \( n_3 \).

**Case III:**

\[
\begin{align*}
\begin{cases}
 v'_1 P < 0 \\
 v'_3 P > 0
\end{cases}
\end{align*}
\]

42
Therefore, the four signs of $\mathbf{n}_j^TP_j$ in this situation will be $(+, -, -, +)$. And two normal vectors are the candidates of the final estimation solution: $\mathbf{n}_1$ and $\mathbf{n}_4$.

Case IV:

\[
\begin{cases}
\mathbf{v}_i^TP < 0 \\
\mathbf{v}_3^TP < 0 \\
\mathbf{v}_4^TP > 0 \\
\mathbf{v}_5^TP > 0 
\end{cases}
\]

\[
\begin{aligned}
\mathbf{n}_1^TP &= \frac{(1-\lambda_3)(\mathbf{v}_3^TP)^2}{\lambda_1-\lambda_3} - \frac{(\lambda_1-1)(\mathbf{v}_1^TP)^2}{\lambda_1-\lambda_3} > 0 \\
\mathbf{n}_2^TP &= \frac{(1-\lambda_3)(\mathbf{v}_3^TP)^2}{\lambda_1-\lambda_3} + \frac{(\lambda_1-1)(\mathbf{v}_1^TP)^2}{\lambda_1-\lambda_3} < 0 \\
\mathbf{n}_3^TP &= -\frac{(1-\lambda_3)(\mathbf{v}_3^TP)^2}{\lambda_1-\lambda_3} - \frac{(\lambda_1-1)(\mathbf{v}_1^TP)^2}{\lambda_1-\lambda_3} < 0 \\
\mathbf{n}_4^TP &= -\frac{(1-\lambda_3)(\mathbf{v}_3^TP)^2}{\lambda_1-\lambda_3} + \frac{(\lambda_1-1)(\mathbf{v}_1^TP)^2}{\lambda_1-\lambda_3} > 0
\end{aligned}
\]

(3-18)

Therefore, the four signs of $\mathbf{n}_j^TP_j$ in this situation will be $(-, +, +, -)$. And two normal vectors are the candidates of the final estimation solution: $\mathbf{n}_2$ and $\mathbf{n}_3$.

Second, when $\mathbf{P}'\mathbf{P} < \mathbf{Q}'\mathbf{Q}$, Eq. (3-14) will be the constraint:

Constraint 2: $(1-\lambda_3)(\mathbf{v}_3^TP)^2 < (\lambda_1-1)(\mathbf{v}_1^TP)^2$

Case I:

\[
\begin{cases}
\mathbf{v}_i^TP > 0 \\
\mathbf{v}_3^TP > 0 
\end{cases}
\]
\[
\begin{align*}
n_1'P &= (1 - \lambda_3)(v_1'P)^2 + (\lambda_1 - 1)(v_1'P)^2 > 0 \\
n_2'P &= -\sqrt{\frac{(1 - \lambda_3)(v_2'P)^2}{\lambda_1 - \lambda_3}} - \sqrt{\frac{(\lambda_1 - 1)(v_2'P)^2}{\lambda_1 - \lambda_3}} < 0 \\
n_3'P &= -\sqrt{\frac{(1 - \lambda_3)(v_3'P)^2}{\lambda_1 - \lambda_3}} + \sqrt{\frac{(\lambda_1 - 1)(v_3'P)^2}{\lambda_1 - \lambda_3}} > 0 \\
n_4'P &= -\sqrt{\frac{(1 - \lambda_3)(v_4'P)^2}{\lambda_1 - \lambda_3}} - \sqrt{\frac{(\lambda_1 - 1)(v_4'P)^2}{\lambda_1 - \lambda_3}} < 0
\end{align*}
\] (3-19)

Therefore, the four signs of \( n_j'P \) in this situation will be \(+, -, +, -\). And two normal vectors are the candidates of the final estimation solution: \( n_1 \) and \( n_3 \).

Case II: \[
\begin{align*}
v_1'P &= > 0 \\
v_3'P &= < 0
\end{align*}
\]

\[
\begin{align*}
n_1'P &= -\sqrt{\frac{(1 - \lambda_3)(v_1'P)^2}{\lambda_1 - \lambda_3}} + \sqrt{\frac{(\lambda_1 - 1)(v_1'P)^2}{\lambda_1 - \lambda_3}} > 0 \\
n_2'P &= -\sqrt{\frac{(1 - \lambda_3)(v_2'P)^2}{\lambda_1 - \lambda_3}} - \sqrt{\frac{(\lambda_1 - 1)(v_2'P)^2}{\lambda_1 - \lambda_3}} < 0 \\
n_3'P &= -\sqrt{\frac{(1 - \lambda_3)(v_3'P)^2}{\lambda_1 - \lambda_3}} + \sqrt{\frac{(\lambda_1 - 1)(v_3'P)^2}{\lambda_1 - \lambda_3}} > 0 \\
n_4'P &= -\sqrt{\frac{(1 - \lambda_3)(v_4'P)^2}{\lambda_1 - \lambda_3}} - \sqrt{\frac{(\lambda_1 - 1)(v_4'P)^2}{\lambda_1 - \lambda_3}} < 0
\end{align*}
\] (3-20)

Therefore, the four signs of \( n_j'P \) in this situation will be \(+, -, +, -\). And two normal vectors are the candidates of the final estimation solution: \( n_1 \) and \( n_3 \).

Case III: \[
\begin{align*}
v_1'P &= < 0 \\
v_3'P &= > 0
\end{align*}
\]
Therefore, the four signs of \( \mathbf{n}_j^	op \mathbf{P}_i \) in this situation will be \((-,-,+,+)\). And two normal vectors are the candidates of the final estimation solution: \( \mathbf{n}_2 \) and \( \mathbf{n}_4 \).

\[
\begin{align*}
\mathbf{n}_1' \mathbf{P} &= -\sqrt{\frac{(1-\lambda_3)(\mathbf{v}_3' \mathbf{P})^2}{\lambda_4 - \lambda_3}} - \sqrt{\frac{(\lambda_4 - 1)(\mathbf{v}_3' \mathbf{P})^2}{\lambda_4 - \lambda_3}} < 0 \\
\mathbf{n}_2' \mathbf{P} &= -\sqrt{\frac{(1-\lambda_3)(\mathbf{v}_3' \mathbf{P})^2}{\lambda_4 - \lambda_3}} + \sqrt{\frac{(\lambda_4 - 1)(\mathbf{v}_3' \mathbf{P})^2}{\lambda_4 - \lambda_3}} > 0 \\
\mathbf{n}_3' \mathbf{P} &= -\sqrt{\frac{(1-\lambda_3)(\mathbf{v}_3' \mathbf{P})^2}{\lambda_4 - \lambda_3}} - \sqrt{\frac{(\lambda_4 - 1)(\mathbf{v}_3' \mathbf{P})^2}{\lambda_4 - \lambda_3}} < 0 \\
\mathbf{n}_4' \mathbf{P} &= -\sqrt{\frac{(1-\lambda_3)(\mathbf{v}_3' \mathbf{P})^2}{\lambda_4 - \lambda_3}} + \sqrt{\frac{(\lambda_4 - 1)(\mathbf{v}_3' \mathbf{P})^2}{\lambda_4 - \lambda_3}} > 0 \\
\end{align*}
\]

Therefore, the four signs of \( \mathbf{n}_j' \mathbf{P}_i \) in this situation will be \((-,-,+,+)\). And two normal vectors are the candidates of the final estimation solution: \( \mathbf{n}_2 \) and \( \mathbf{n}_4 \).

Finally, it is possible to determine the sign of \( \mathbf{n}_j' \mathbf{P}_i \) for each possible case (\( \text{I}, \text{II}, \text{III}, \) and \( \text{IV} \)), as listed in Table 2. That is, there are four cases for \( \mathbf{P}' \mathbf{P} > \mathbf{Q}' \mathbf{Q} \), and four cases for \( \mathbf{P}' \mathbf{P} < \mathbf{Q}' \mathbf{Q} \). If the signs of \( \mathbf{n}_j' \mathbf{P}_i \) for all the corresponding feature points are positive, \( \mathbf{n}_j \) is one of the solutions.
<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P'P &gt; Q'Q$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n_i'P$</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>$n_j'P$</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>$n_j'P$</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>$n_i'P$</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>$P'P &lt; Q'Q$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n_i'P$</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$n_j'P$</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>$n_j'P$</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$n_i'P$</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

By using Table 2, the sign of Eq. (3-6) for all possible situations are quite clear at a glance. However, considering the reality of the 3D movement of a rigid object, all of the eight cases above (in Table 2) cannot be arbitrarily combined in a real 3D movement. Some other constraints are necessary.

Basically, during 3D movement of a rigid object, the distance between any two feature points on the same feature plane does not change. This property can be expressed as

$$\|P_i - P_j\| = \|Q_i - Q_j\| \quad (i \neq j)$$

$$\begin{pmatrix} P_i - P_j \end{pmatrix} \begin{pmatrix} P_i - P_j \end{pmatrix} = \begin{pmatrix} Q_i - Q_j \end{pmatrix} \begin{pmatrix} Q_i - Q_j \end{pmatrix}$$

(3-23)

Considering the homography relation between $P_i$ and $Q_i$, Eq. (3-8), $Q_i = HP_i$, and the SVD of homography matrix $H$, Eq. (3-9), $H = U \Sigma V'$, the right-hand side of Eq.
Another way, because right singular vector matrix \( \textbf{V} \) is the orthogonal matrix, the norm of \( (\textbf{P}_i - \textbf{P}_j) \) does not change before and after \( \textbf{P}_i \) and \( \textbf{P}_j \) is rotated by \( \textbf{V} \).

Therefore, the left-hand side of Eq. (3-23) can be rewritten as

\[
(\textbf{P}_i - \textbf{P}_j)(\textbf{P}_i - \textbf{P}_j) = (\textbf{V}'\textbf{P}_i - \textbf{V}'\textbf{P}_j)(\textbf{V}'\textbf{P}_i - \textbf{V}'\textbf{P}_j)
\]

\[
= (\nu_i^{\prime 2}_i \textbf{P}_i - \nu_i^{\prime 2}_i \textbf{P}_j)^2 + (\nu_j^{\prime 2}_i \textbf{P}_i - \nu_j^{\prime 2}_i \textbf{P}_j)^2 + (\nu_j^{\prime 2}_j \textbf{P}_i - \nu_j^{\prime 2}_j \textbf{P}_j)^2
\]

Substituting Eq. (3-24) and Eq. (3-25) into Eq. (3-23), it is possible to obtain the constraint for corresponding feature points, which must always be satisfied in real 3D movement by a rigid object:

\[
(\lambda_i - 1)(\nu_i^{\prime 2}_i \textbf{P}_i - \nu_j^{\prime 2}_i \textbf{P}_j)^2 = (1 - \lambda_j)(\nu_j^{\prime 2}_i \textbf{P}_i - \nu_j^{\prime 2}_j \textbf{P}_j)^2
\]

\[\Rightarrow (\lambda_i - 1)(\nu_i^{\prime 2}_i \textbf{P}_i)^2 + (\lambda_i - 1)(\nu_j^{\prime 2}_i \textbf{P}_i)^2 - 2(\lambda_i - 1)(\nu_i^{\prime 2}_i \textbf{P}_i)(\nu_j^{\prime 2}_i \textbf{P}_j) \]

\[= (1 - \lambda_j)(\nu_j^{\prime 2}_i \textbf{P}_i)^2 + (1 - \lambda_j)(\nu_j^{\prime 2}_j \textbf{P}_j)^2 - 2(1 - \lambda_j)(\nu_j^{\prime 2}_i \textbf{P}_i)(\nu_j^{\prime 2}_j \textbf{P}_j)
\]

Eq. (3-26) is the constraint of “rigid object 3D movement reality”.

If \( \forall i, \textbf{P}_i^\prime > \textbf{Q}_i Q_i \), considering Eq. (3-13), \( (1 - \lambda_j)(\nu_j^{\prime 2}_i \textbf{P})^2 > (\lambda_i - 1)(\nu_i^{\prime 2}_i \textbf{P})^2 \). In order to satisfy the “rigid object 3D movement reality”, i.e. make Eq. (3-26) established, \( \nu_j^{\prime 2}_i \textbf{P}_i \) and \( \nu_j^{\prime 2}_j \textbf{P}_j \) must not have different signs (no matter whether \( \nu_i^{\prime 2}_i \textbf{P}_i \) and \( \nu_i^{\prime 2}_j \textbf{P}_j \) have different signs or not). Because

\[
\therefore (1 - \lambda_j)(\nu_j^{\prime 2}_i \textbf{P})^2 > (\lambda_i - 1)(\nu_i^{\prime 2}_i \textbf{P})^2
\]

\[
\therefore (1 - \lambda_j)(\nu_j^{\prime 2}_j \textbf{P})^2 > (\lambda_i - 1)(\nu_i^{\prime 2}_j \textbf{P})^2
\]

\[
(1 - \lambda_j)(\nu_j^{\prime 2}_i \textbf{P})^2 > (\lambda_i - 1)(\nu_i^{\prime 2}_i \textbf{P})^2 > (\lambda_i - 1)(\nu_i^{\prime 2}_j \textbf{P})^2
\]

\[
\therefore (1 - \lambda_j)(\nu_j^{\prime 2}_i \textbf{P})^2 + (1 - \lambda_j)(\nu_j^{\prime 2}_j \textbf{P})^2 > (\lambda_i - 1)(\nu_i^{\prime 2}_i \textbf{P})^2 + (\lambda_i - 1)(\nu_i^{\prime 2}_j \textbf{P})^2
\]
To establish Eq. (3-26), there should be

$$-2(\lambda_1 - 1)(v_i^T P_i)(v_j^T P_j) > -2(1 - \lambda_3)(v_i^T P_i)(v_j^T P_j)$$

(3-28)

Considering Eq. (3-28), if $v_i^T P_i$ and $v_j^T P_j$ have the different signs, $-2(1 - \lambda_3)(v_i^T P_i)(v_j^T P_j)$ will be positive, the right side of Eq. (3-26) will always larger than the left side, whatever $-2(\lambda_1 - 1)(v_i^T P_i)(v_j^T P_j)$ is. Therefore, “$v_i^T P_i$ and $v_j^T P_j$ have the same signs” is a necessary condition to establish Eq. (3-26) for $\forall i$, $P_i^T P_i > Q_i^T Q_i$.

Similarly, if $\forall i$, $P_i^T P_i < Q_i^T Q_i$, considering Eq. (3-14), $(1 - \lambda_3)(v_i^T P_i)^2 < (\lambda_1 - 1)(v_i^T P_i)^2$. In order to establish Eq. (3-26), $v_i^T P_i$ and $v_i^T P_j$ must not have different signs (no matter whether $v_i^T P_i$ and $v_j^T P_j$ have different signs or not). Because

$$\therefore (1 - \lambda_3)(v_i^T P_i)^2 < (\lambda_1 - 1)(v_i^T P_i)^2$$
$$\therefore (1 - \lambda_3)(v_i^T P_i)^2 < (\lambda_1 - 1)(v_i^T P_i)^2$$
$$\therefore (1 - \lambda_3)(v_i^T P_i)^2 + (1 - \lambda_3)(v_i^T P_j)^2 < (\lambda_1 - 1)(v_i^T P_i)^2 + (\lambda_1 - 1)(v_i^T P_j)^2$$

(3-29)

To establish Eq. (3-26), there should be

$$-2(\lambda_1 - 1)(v_i^T P_i)(v_j^T P_j) < -2(1 - \lambda_3)(v_i^T P_i)(v_j^T P_j)$$

(3-30)

Considering Eq. (3-30), if $v_i^T P_i$ and $v_j^T P_j$ have the different signs, $-2(1 - \lambda_3)(v_i^T P_i)(v_j^T P_j)$ will be positive, the left side of Eq. (3-26) will always larger than the right side, whatever $-2(\lambda_1 - 1)(v_i^T P_i)(v_j^T P_j)$ is. Therefore, “$v_i^T P_i$ and $v_j^T P_j$ have the same signs” is a necessary condition to establish Eq. (3-26) for $\forall i$, $P_i^T P_i < Q_i^T Q_i$.

Recalling Table 1 and considering what has been discussed above, it is obvious that:

- For $\forall i$, $P_i^T P_i > Q_i^T Q_i$, only I and III (or II and IV) in Table 1 will occur simultaneously.
For $\forall i$, $P_i^jP_j < Q_i^jQ_j$, only $\text{I}$ and $\text{II}$ (or $\text{III}$ and $\text{IV}$) in Table 1 will occur simultaneously.

On the other hand, it is possible to group the feature points based on the two distance relations of Eq. (3-7), $P_i^jP_j > Q_i^jQ_j$ and $P_i^jP_j < Q_i^jQ_j$. No matter how many feature points there are, if they are grouped two by two, they will always belong to one of the three kinds of possibilities below:

1. Both of two points satisfy $P_i^jP_j > Q_i^jQ_j$
2. Both of two points satisfy $P_i^jP_j < Q_i^jQ_j$
3. $P_i^jP_j > Q_i^jQ_j$ and $P_j^iP_i < Q_j^iQ_j$ ($i \neq j$).

Now let discuss the actual estimated solutions case by case, based on Table 2.

**Possibility (1): Both of two points satisfy $P_i^jP_j > Q_i^jQ_j$**

Note: Only $\text{I}$ and $\text{III}$ (or $\text{II}$ and $\text{IV}$) in Table 1 will occur simultaneously in this situation.

1. $P_i \in \text{Case I}$, $P_j \in \text{Case I}$

   \[
   \begin{align*}
   n_i^iP_i > 0, & \quad n_j^iP_j > 0 \quad \Rightarrow n_1 \\
   n_i^iP_i < 0, & \quad n_j^iP_j < 0 \quad \Rightarrow \times \\
   n_i^iP_i < 0, & \quad n_j^iP_j < 0 \quad \Rightarrow \times \\
   n_i^iP_i > 0, & \quad n_j^iP_j > 0 \quad \Rightarrow n_4
   \end{align*}
   \] (3-31)

   In this case, $n_1$ and $n_4$ are the estimated solutions.

2. $P_i \in \text{Case I}$, $P_j \in \text{Case III}$

   \[
   \begin{align*}
   n_i^iP_i > 0, & \quad n_j^iP_j > 0 \quad \Rightarrow n_1 \\
   n_i^iP_i < 0, & \quad n_j^iP_j < 0 \quad \Rightarrow \times \\
   n_i^iP_i < 0, & \quad n_j^iP_j < 0 \quad \Rightarrow \times \\
   n_i^iP_i > 0, & \quad n_j^iP_j > 0 \quad \Rightarrow n_4
   \end{align*}
   \] (3-32)

   In this case, $n_1$ and $n_4$ are the estimated solutions.

3. $P_i \in \text{Case III}$, $P_j \in \text{Case I}$
\[
\begin{align*}
&\begin{cases}
  n'_i P_i > 0, & n'_j P_j > 0 \quad \Rightarrow \quad n_1 \\
n'_i P_i < 0, & n'_j P_j < 0 \quad \Rightarrow \quad \times \\
n'_i P_i < 0, & n'_j P_j < 0 \quad \Rightarrow \quad \times \\
n'_i P_i > 0, & n'_j P_j > 0 \quad \Rightarrow \quad n_4 
\end{cases} \\
\end{align*}
\]

In this case, \( n_1 \) and \( n_4 \) are the estimated solutions.

\( (4) \ P_i \in \text{Case III}, \quad P_j \in \text{Case III} \)
\[
\begin{align*}
&\begin{cases}
  n'_i P_i > 0, & n'_j P_j > 0 \quad \Rightarrow \quad n_1 \\
n'_i P_i < 0, & n'_j P_j < 0 \quad \Rightarrow \quad \times \\
n'_i P_i < 0, & n'_j P_j < 0 \quad \Rightarrow \quad \times \\
n'_i P_i > 0, & n'_j P_j > 0 \quad \Rightarrow \quad n_4 
\end{cases} \\
\end{align*}
\]

In this case, \( n_1 \) and \( n_4 \) are the estimated solutions.

\( (5) \ P_i \in \text{Case II}, \quad P_j \in \text{Case II} \)
\[
\begin{align*}
&\begin{cases}
  n'_i P_i < 0, & n'_j P_j < 0 \quad \Rightarrow \quad \times \\
n'_i P_i > 0, & n'_j P_j > 0 \quad \Rightarrow \quad n_2 \\
n'_i P_i > 0, & n'_j P_j > 0 \quad \Rightarrow \quad n_3 \\
n'_i P_i < 0, & n'_j P_j < 0 \quad \Rightarrow \quad \times 
\end{cases} \\
\end{align*}
\]

In this case, \( n_2 \) and \( n_3 \) are the estimated solutions.

\( (6) \ P_i \in \text{Case II}, \quad P_j \in \text{Case IV} \)
\[
\begin{align*}
&\begin{cases}
  n'_i P_i < 0, & n'_j P_j < 0 \quad \Rightarrow \quad \times \\
n'_i P_i > 0, & n'_j P_j > 0 \quad \Rightarrow \quad n_2 \\
n'_i P_i > 0, & n'_j P_j > 0 \quad \Rightarrow \quad n_3 \\
n'_i P_i < 0, & n'_j P_j < 0 \quad \Rightarrow \quad \times 
\end{cases} \\
\end{align*}
\]

In this case, \( n_2 \) and \( n_3 \) are the estimated solutions.

\( (7) \ P_i \in \text{Case IV}, \quad P_j \in \text{Case II} \)
\[
\begin{align*}
&\begin{cases}
  n'_i P_i < 0, & n'_j P_j < 0 \quad \Rightarrow \quad \times \\
n'_i P_i > 0, & n'_j P_j > 0 \quad \Rightarrow \quad n_2 \\
n'_i P_i > 0, & n'_j P_j > 0 \quad \Rightarrow \quad n_3 \\
n'_i P_i < 0, & n'_j P_j < 0 \quad \Rightarrow \quad \times 
\end{cases} \\
\end{align*}
\]

In this case, \( n_2 \) and \( n_3 \) are the estimated solutions.
(8) \( P_i \in \text{Case IV}, \ P_j \in \text{Case IV} \)

\[
\begin{aligned}
&n_1'P_i < 0, \ n_1'P_j < 0 \Rightarrow \times \\
&n_2'P_i > 0, \ n_2'P_j > 0 \Rightarrow n_2 \\
&n_3'P_i > 0, \ n_3'P_j > 0 \Rightarrow n_3 \\
&n_4'P_i < 0, \ n_4'P_j < 0 \Rightarrow \times
\end{aligned}
\] (3-38)

In this case, \( n_2 \) and \( n_3 \) are the estimated solutions.

**Possibility (2): Both of two points satisfy \( P'P < Q'Q \)**

Note: Only I and II (or III and IV) in Table 1 will occur simultaneously in this situation.

(1) \( P_i \in \text{Case I}, \ P_j \in \text{Case I} \)

\[
\begin{aligned}
&n_1'P_i > 0, \ n_1'P_j > 0 \Rightarrow n_1 \\
&n_2'P_i < 0, \ n_2'P_j < 0 \Rightarrow \times \\
&n_3'P_i > 0, \ n_3'P_j > 0 \Rightarrow n_3 \\
&n_4'P_i < 0, \ n_4'P_j < 0 \Rightarrow \times
\end{aligned}
\] (3-39)

In this case, \( n_1 \) and \( n_3 \) are the estimated solutions.

(2) \( P_i \in \text{Case I}, \ P_j \in \text{Case II} \)

\[
\begin{aligned}
&n_1'P_i > 0, \ n_1'P_j > 0 \Rightarrow n_1 \\
&n_2'P_i < 0, \ n_2'P_j < 0 \Rightarrow \times \\
&n_3'P_i > 0, \ n_3'P_j > 0 \Rightarrow n_3 \\
&n_4'P_i < 0, \ n_4'P_j < 0 \Rightarrow \times
\end{aligned}
\] (3-40)

In this case, \( n_1 \) and \( n_3 \) are the estimated solutions.

(3) \( P_i \in \text{Case II}, \ P_j \in \text{Case I} \)

\[
\begin{aligned}
&n_1'P_i > 0, \ n_1'P_j > 0 \Rightarrow n_1 \\
&n_2'P_i < 0, \ n_2'P_j < 0 \Rightarrow \times \\
&n_3'P_i > 0, \ n_3'P_j > 0 \Rightarrow n_3 \\
&n_4'P_i < 0, \ n_4'P_j < 0 \Rightarrow \times
\end{aligned}
\] (3-41)

In this case, \( n_1 \) and \( n_3 \) are the estimated solutions.
(4) $P_i \in \text{Case II}, \quad P_j \in \text{Case II}$
\[
\begin{align*}
&n'_i P_i > 0, \quad n'_j P_j > 0 \Rightarrow n_1 \\
&n'_i P_i < 0, \quad n'_j P_j < 0 \Rightarrow \times \\
&n'_i P_i > 0, \quad n'_j P_j > 0 \Rightarrow n_3 \\
&n'_i P_i < 0, \quad n'_j P_j < 0 \Rightarrow \times \\
\end{align*}
\]
(3-42)

In this case, $n_1$ and $n_3$ are the estimated solutions.

(5) $P_i \in \text{Case III}, \quad P_j \in \text{Case III}$
\[
\begin{align*}
&n'_i P_i < 0, \quad n'_j P_j < 0 \Rightarrow \times \\
&n'_i P_i > 0, \quad n'_j P_j > 0 \Rightarrow n_2 \\
&n'_i P_i < 0, \quad n'_j P_j < 0 \Rightarrow \times \\
&n'_i P_i > 0, \quad n'_j P_j > 0 \Rightarrow n_4 \\
\end{align*}
\]
(3-43)

In this case, $n_1$ and $n_3$ are the estimated solutions.

(6) $P_i \in \text{Case III}, \quad P_j \in \text{Case IV}$
\[
\begin{align*}
&n'_i P_i < 0, \quad n'_j P_j < 0 \Rightarrow \times \\
&n'_i P_i > 0, \quad n'_j P_j > 0 \Rightarrow n_2 \\
&n'_i P_i < 0, \quad n'_j P_j < 0 \Rightarrow \times \\
&n'_i P_i > 0, \quad n'_j P_j > 0 \Rightarrow n_4 \\
\end{align*}
\]
(3-44)

In this case, $n_1$ and $n_3$ are the estimated solutions.

(7) $P_i \in \text{Case IV}, \quad P_j \in \text{Case III}$
\[
\begin{align*}
&n'_i P_i < 0, \quad n'_j P_j < 0 \Rightarrow \times \\
&n'_i P_i > 0, \quad n'_j P_j > 0 \Rightarrow n_2 \\
&n'_i P_i < 0, \quad n'_j P_j < 0 \Rightarrow \times \\
&n'_i P_i > 0, \quad n'_j P_j > 0 \Rightarrow n_4 \\
\end{align*}
\]
(3-45)

In this case, $n_1$ and $n_3$ are the estimated solutions.

(8) $P_i \in \text{Case IV}, \quad P_j \in \text{Case IV}$
\[
\begin{align*}
&n'_i P_i < 0, \quad n'_j P_j < 0 \Rightarrow \times \\
&n'_i P_i > 0, \quad n'_j P_j > 0 \Rightarrow n_2 \\
&n'_i P_i < 0, \quad n'_j P_j < 0 \Rightarrow \times \\
&n'_i P_i > 0, \quad n'_j P_j > 0 \Rightarrow n_4 \\
\end{align*}
\]
(3-46)
In this case, \( \textbf{n}_1 \) and \( \textbf{n}_3 \) are the estimated solutions.

As a conclusion, Table 3 summarizes the solutions to all combinations for possibility (1) or (2) which are discussed above. Cases \( \text{I}, \text{II}, \text{III}, \text{and IV} \) on the horizontal and vertical axes are listed in Table 1. Normal vectors \( \textbf{n}_j \) that represent the solutions to all combinations, can be determined by combining two of the eight cases in Table 2.

Table 3 Solutions to possibility (1) or (2)

<table>
<thead>
<tr>
<th></th>
<th>( \textbf{P}'\textbf{P} ) &gt; ( \textbf{Q}'\textbf{Q} )</th>
<th>( \textbf{P}'\textbf{P} ) &lt; ( \textbf{Q}'\textbf{Q} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \textbf{P}_1 )</td>
<td>( \textbf{P}_2 )</td>
</tr>
<tr>
<td>( \textbf{P}_1 )</td>
<td>( \textbf{n}_1, \textbf{n}_4 )</td>
<td>( \textbf{n}_1, \textbf{n}_4 )</td>
</tr>
<tr>
<td>( \textbf{P}_2 )</td>
<td>( \textbf{n}_2, \textbf{n}_3 )</td>
<td>( \textbf{n}_2, \textbf{n}_3 )</td>
</tr>
<tr>
<td>( \textbf{P}_3 )</td>
<td>( \textbf{n}_1, \textbf{n}_4 )</td>
<td>( \textbf{n}_1, \textbf{n}_4 )</td>
</tr>
<tr>
<td>( \textbf{P}_4 )</td>
<td>( \textbf{n}_2, \textbf{n}_3 )</td>
<td>( \textbf{n}_2, \textbf{n}_3 )</td>
</tr>
</tbody>
</table>

The blank cells in Table 3 mean no solution. Based on the reality constraint of rigid object 3D motion discussed above, these combinations will not occur in real 3D
motion. Thus, they can be ignored. The others obtain ambiguous solutions. Therefore, if all the feature points belong to possibility (1) or (2), solution ambiguity will be inevitable.

Possibility (3): $P_i^TP_i > Q_i^TQ_i$ and $P_j^TP_j < Q_j^TQ_j$ ($i \neq j$).

Note: the “rigid object 3D movement reality” is not available in this situation, so $P_i$ and $P_j$ can be arbitrarily combined of all four cases in Table 1.

(1) $P_i \in \text{Case I}, \ P_j \in \text{Case I}$

\[
\begin{align*}
&n_1^TP_i > 0, \ n_1^TP_j > 0 \quad \Rightarrow \quad n_1 \\
&n_1^TP_i < 0, \ n_1^TP_j < 0 \quad \Rightarrow \quad \times \\
&n_1^TP_i < 0, \ n_1^TP_j > 0 \quad \Rightarrow \quad \times \\
&n_1^TP_i > 0, \ n_1^TP_j < 0 \quad \Rightarrow \quad \times
\end{align*}
\]

(3-47)

In this case, only $n_1$ are the estimated solutions.

(2) $P_i \in \text{Case I}, \ P_j \in \text{Case II}$

\[
\begin{align*}
&n_1^TP_i > 0, \ n_2^TP_j > 0 \quad \Rightarrow \quad n_1 \\
&n_2^TP_i < 0, \ n_2^TP_j < 0 \quad \Rightarrow \quad \times \\
&n_2^TP_i < 0, \ n_2^TP_j > 0 \quad \Rightarrow \quad \times \\
&n_2^TP_i > 0, \ n_2^TP_j < 0 \quad \Rightarrow \quad \times
\end{align*}
\]

(3-48)

In this case, only $n_1$ are the estimated solutions.

(3) $P_i \in \text{Case I}, \ P_j \in \text{Case III}$

\[
\begin{align*}
&n_1^TP_i > 0, \ n_3^TP_j < 0 \quad \Rightarrow \quad \times \\
&n_3^TP_i < 0, \ n_3^TP_j > 0 \quad \Rightarrow \quad \times \\
&n_3^TP_i < 0, \ n_3^TP_j < 0 \quad \Rightarrow \quad \times \\
&n_3^TP_i > 0, \ n_3^TP_j > 0 \quad \Rightarrow \quad n_4
\end{align*}
\]

(3-49)

In this case, only $n_4$ are the estimated solutions.

(4) $P_i \in \text{Case I}, \ P_j \in \text{Case IV}$
\[
\begin{aligned}
&\left\{ \begin{array}{l}
n'_i P_i > 0, \quad n'_j P_j < 0 \quad \Rightarrow \quad \times \\
n'_i P_i < 0, \quad n'_j P_j > 0 \quad \Rightarrow \quad \times
\end{array} \right. \\
&\left\{ \begin{array}{l}
n'_i P_i > 0, \quad n'_j P_j < 0 \quad \Rightarrow \quad \times \\
n'_i P_i < 0, \quad n'_j P_j > 0 \quad \Rightarrow \quad n_4
\end{array} \right.
\end{aligned}
\]

In this case, only \( n_4 \) are the estimated solutions.

(5) \( P_i \in \text{Case II}, \ P_j \in \text{Case I} \)

\[
\begin{aligned}
&\left\{ \begin{array}{l}
n'_i P_i < 0, \quad n'_j P_j > 0 \quad \Rightarrow \quad \times \\
n'_i P_i > 0, \quad n'_j P_j < 0 \quad \Rightarrow \quad \times
\end{array} \right. \\
&\left\{ \begin{array}{l}
n'_i P_i > 0, \quad n'_j P_j > 0 \quad \Rightarrow \quad n_3 \\
n'_i P_i < 0, \quad n'_j P_j < 0 \quad \Rightarrow \quad \times
\end{array} \right.
\end{aligned}
\]

In this case, only \( n_3 \) are the estimated solutions.

(6) \( P_i \in \text{Case II}, \ P_j \in \text{Case II} \)

\[
\begin{aligned}
&\left\{ \begin{array}{l}
n'_i P_i < 0, \quad n'_j P_j > 0 \quad \Rightarrow \quad \times \\
n'_i P_i > 0, \quad n'_j P_j < 0 \quad \Rightarrow \quad \times
\end{array} \right. \\
&\left\{ \begin{array}{l}
n'_i P_i > 0, \quad n'_j P_j > 0 \quad \Rightarrow \quad n_3 \\
n'_i P_i < 0, \quad n'_j P_j < 0 \quad \Rightarrow \quad \times
\end{array} \right.
\end{aligned}
\]

In this case, only \( n_3 \) are the estimated solutions.

(7) \( P_i \in \text{Case II}, \ P_j \in \text{Case III} \)

\[
\begin{aligned}
&\left\{ \begin{array}{l}
n'_i P_i < 0, \quad n'_j P_j < 0 \quad \Rightarrow \quad \times \\
n'_i P_i > 0, \quad n'_j P_j > 0 \quad \Rightarrow \quad n_2 \\
n'_i P_i > 0, \quad n'_j P_j < 0 \quad \Rightarrow \quad \times \\
n'_i P_i < 0, \quad n'_j P_j > 0 \quad \Rightarrow \quad \times
\end{array} \right.
\end{aligned}
\]

In this case, only \( n_2 \) are the estimated solutions.

(8) \( P_i \in \text{Case II}, \ P_j \in \text{Case IV} \)

\[
\begin{aligned}
&\left\{ \begin{array}{l}
n'_i P_i < 0, \quad n'_j P_j < 0 \quad \Rightarrow \quad \times \\
n'_i P_i > 0, \quad n'_j P_j > 0 \quad \Rightarrow \quad n_2 \\
n'_i P_i > 0, \quad n'_j P_j < 0 \quad \Rightarrow \quad \times \\
n'_i P_i < 0, \quad n'_j P_j > 0 \quad \Rightarrow \quad \times
\end{array} \right.
\end{aligned}
\]

In this case, only \( n_2 \) are the estimated solutions.
(9) $P_i \in \text{Case III}, \ P_j \in \text{Case I}$

\[
\begin{align*}
&n'_i P_i > 0, \ n'_j P_j > 0 \Rightarrow n_1 \\
n'_i P_i < 0, \ n'_j P_j < 0 \Rightarrow \times \\
n'_i P_i < 0, \ n'_j P_j > 0 \Rightarrow \times \\
n'_i P_i > 0, \ n'_j P_j < 0 \Rightarrow \times 
\end{align*}
\]

(3-55)

In this case, only $n_1$ are the estimated solutions.

(10) $P_i \in \text{Case III}, \ P_j \in \text{Case II}$

\[
\begin{align*}
&n'_i P_i > 0, \ n'_j P_j > 0 \Rightarrow n_1 \\
n'_i P_i < 0, \ n'_j P_j < 0 \Rightarrow \times \\
n'_i P_i < 0, \ n'_j P_j > 0 \Rightarrow \times \\
n'_i P_i > 0, \ n'_j P_j < 0 \Rightarrow \times 
\end{align*}
\]

(3-56)

In this case, only $n_1$ are the estimated solutions.

(11) $P_i \in \text{Case III}, \ P_j \in \text{Case III}$

\[
\begin{align*}
&n'_i P_i > 0, \ n'_j P_j < 0 \Rightarrow \times \\
n'_i P_i < 0, \ n'_j P_j > 0 \Rightarrow \times \\
n'_i P_i < 0, \ n'_j P_j < 0 \Rightarrow \times \\
n'_i P_i > 0, \ n'_j P_j > 0 \Rightarrow n_4 
\end{align*}
\]

(3-57)

In this case, only $n_4$ are the estimated solutions.

(12) $P_i \in \text{Case III}, \ P_j \in \text{Case IV}$

\[
\begin{align*}
&n'_i P_i > 0, \ n'_j P_j < 0 \Rightarrow \times \\
n'_i P_i < 0, \ n'_j P_j > 0 \Rightarrow \times \\
n'_i P_i < 0, \ n'_j P_j < 0 \Rightarrow \times \\
n'_i P_i > 0, \ n'_j P_j > 0 \Rightarrow n_4 
\end{align*}
\]

(3-58)

In this case, only $n_4$ are the estimated solutions.

(13) $P_i \in \text{Case IV}, \ P_j \in \text{Case I}$

56
\[
\begin{align*}
&\begin{cases} 
n_i'p_i < 0, & n_j'p_j > 0 \Rightarrow \times \\
n_j'p_j > 0, & n_j'p_j < 0 \Rightarrow \times \\
n_i'p_i > 0, & n_j'p_j > 0 \Rightarrow n_3 \\
n_i'p_i > 0, & n_j'p_j < 0 \Rightarrow n_3 \\
n_i'p_i < 0, & n_j'p_j < 0 \Rightarrow \times 
\end{cases} 
\end{align*}
\] (3-59)

In this case, only \( n_3 \) are the estimated solutions.

(14) \( P_i \in \text{Case IV}, \quad P_j \in \text{Case II} \)
\[
\begin{align*}
&\begin{cases} 
n_i'p_i < 0, & n_j'p_j > 0 \Rightarrow \times \\
n_j'p_j > 0, & n_j'p_j < 0 \Rightarrow \times \\
n_i'p_i > 0, & n_j'p_j > 0 \Rightarrow n_3 \\
n_i'p_i > 0, & n_j'p_j < 0 \Rightarrow \times \\
n_i'p_i < 0, & n_j'p_j < 0 \Rightarrow \times 
\end{cases} 
\end{align*}
\] (3-60)

In this case, only \( n_3 \) are the estimated solutions.

(15) \( P_i \in \text{Case IV}, \quad P_j \in \text{Case III} \)
\[
\begin{align*}
&\begin{cases} 
n_i'p_i < 0, & n_j'p_j < 0 \Rightarrow \times \\
n_j'p_j > 0, & n_j'p_j > 0 \Rightarrow n_2 \\
n_i'p_i > 0, & n_j'p_j < 0 \Rightarrow \times \\
n_i'p_i < 0, & n_j'p_j > 0 \Rightarrow \times 
\end{cases} 
\end{align*}
\] (3-61)

In this case, only \( n_2 \) are the estimated solutions.

(16) \( P_i \in \text{Case IV}, \quad P_j \in \text{Case IV} \)
\[
\begin{align*}
&\begin{cases} 
n_i'p_i < 0, & n_j'p_j < 0 \Rightarrow \times \\
n_j'p_j > 0, & n_j'p_j > 0 \Rightarrow n_2 \\
n_i'p_i > 0, & n_j'p_j < 0 \Rightarrow \times \\
n_i'p_i < 0, & n_j'p_j > 0 \Rightarrow \times 
\end{cases} 
\end{align*}
\] (3-62)

In this case, only \( n_2 \) are the estimated solutions.

As a conclusion, Table 4 summarizes the solutions to possibility (3). Cases \text{I , II , III, and IV} for the horizontal axis are possible under the condition of \( P'P > Q'Q \), and \text{I , II , III, and IV} for the vertical axis are possible cases under the condition of \( P'P < Q'Q \). They are all listed in Table 1. As a result, all the possible combinations for possibility (3) will obtain a unique solution.
Table 4 Solutions to possibility (3)

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_j/\bar{P}_j &gt; Q_j/\bar{Q}_j$</td>
<td>I</td>
<td>II</td>
<td>III</td>
<td>IV</td>
</tr>
<tr>
<td>I</td>
<td>$n_1$</td>
<td>$n_3$</td>
<td>$n_1$</td>
<td>$n_3$</td>
</tr>
<tr>
<td>II</td>
<td>$n_1$</td>
<td>$n_3$</td>
<td>$n_1$</td>
<td>$n_3$</td>
</tr>
<tr>
<td>III</td>
<td>$n_4$</td>
<td>$n_2$</td>
<td>$n_4$</td>
<td>$n_2$</td>
</tr>
<tr>
<td>IV</td>
<td>$n_4$</td>
<td>$n_2$</td>
<td>$n_4$</td>
<td>$n_2$</td>
</tr>
</tbody>
</table>

By using Tables 3 and Table 4, it is possible to determine solutions to all combinations of possibilities (1), (2), and (3). From Table 3, it is possible to know that if all distances from the optical center (camera) to the feature points before 3D motion are closer (possibility 1) or further (possibility 2) than those after 3D motion, the ambiguous solutions are inevitable. Therefore, it is necessary to discuss the situations, in which not all distances from the optical center (camera) to the feature points before 3D motion are closer (or further) than those after 3D motion. That means possibilities (1), (2), and (3) will occur simultaneously. Thus, it is necessary to intersect the three kinds of solutions which are obtained from Tables 3 and 4. As a result, Table 5 summarizes all possible case combinations including possibilities (1), (2), and (3) simultaneously.

For example, from the rigid object 3D motion reality constraint in Eq. (3-26), It is clear that: in possibility (1), only I and III may occur in the same real 3D motion; and only I and II may occur in the same real 3D motion in possibility (2). Therefore, in possibility (3), all combinations of \{ I, III \} for $P_i/\bar{P}_i > Q_i/\bar{Q}_i$ and \{ I, II \} for...
\( \mathbf{P}_j \mathbf{P}_j^\top \mathbf{Q}_j^\top \mathbf{Q}_j \) \( \text{will be possible. That is, when } \mathbf{I} \text{ for } \mathbf{P}_j \mathbf{P}_j^\top \mathbf{Q}_j^\top \mathbf{Q}_j \text{ and } \mathbf{I} \text{ for } \mathbf{P}_j \mathbf{P}_j^\top \mathbf{Q}_j^\top \mathbf{Q}_j \text{ occur, only } \mathbf{n}_1 \text{ can be obtained as summarized in Table 4. Referring to Tables 3 and 4, the intersection of three possibilities is operated as}

\[
\{ \mathbf{n}_1, \mathbf{n}_4 \} \cap \{ \mathbf{n}_1, \mathbf{n}_3 \} \cap \{ \mathbf{n}_1 \} = \{ \mathbf{n}_1 \}. \quad (3-63)
\]

<table>
<thead>
<tr>
<th>Possibility (1)</th>
<th>Possibility (2)</th>
<th>Possibility (3)</th>
<th>Final Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>{ \mathbf{I}, \mathbf{II} }: \mathbf{n}_1, \mathbf{n}_3</td>
<td>{ \mathbf{I}, \mathbf{III} }: \mathbf{n}_1</td>
<td>{ \mathbf{I}, \mathbf{II} }, { \mathbf{I}, \mathbf{III} } : \mathbf{n}_1</td>
<td>\mathbf{n}_1</td>
</tr>
<tr>
<td>{ \mathbf{III}, \mathbf{IV} }: \mathbf{n}_2, \mathbf{n}_4</td>
<td>{ \mathbf{I}, \mathbf{III} }, \mathbf{n}_4</td>
<td>{ \mathbf{I}, \mathbf{III} }, \mathbf{n}_4</td>
<td>\mathbf{n}_4</td>
</tr>
<tr>
<td>{ \mathbf{II}, \mathbf{IV} }: \mathbf{n}_2, \mathbf{n}_3</td>
<td>{ \mathbf{I}, \mathbf{II} }, \mathbf{n}_3</td>
<td>{ \mathbf{II}, \mathbf{IV} }, \mathbf{n}_3</td>
<td>\mathbf{n}_3</td>
</tr>
<tr>
<td>{ \mathbf{III}, \mathbf{IV} }: \mathbf{n}_2, \mathbf{n}_4</td>
<td>{ \mathbf{II}, \mathbf{IV} }, \mathbf{n}_2</td>
<td>{ \mathbf{II}, \mathbf{IV} }, \mathbf{n}_2</td>
<td>\mathbf{n}_2</td>
</tr>
</tbody>
</table>

*(note: “×” means all combinations)*

Based on the solutions to all possible common conditions, it is obvious that only one solution can be obtained when the relations between the feature points contain possibility (3). Therefore, this property of the solution ambiguity problem is described as **Dependence 1:**

**Dependence 1**

“If not all the distances between the feature points and the camera become closer after 3D motion, the solution ambiguity problem obtained by homography decomposition will be avoided.”

Because **Dependence 1** is about property of the solution ambiguity problem in
estimating 3D motion parameters by homography decomposition when one rigid object is moving in the space, the feature region on the object will be fixed during the 3D movement. Therefore, Dependence 1 is defined as “**Dependence on 3D Motion Parameters**” for the solution ambiguity problem in estimating 3D motion parameters by homography decomposition, when the feature region size is fixed.

### 3.1.2 Geometrical Proof of Dependence on 3D Motion Parameters

In this section, I will geometrical prove the Dependence 1 which is proposed in last section.

Considering the homography mapping which is illustrated in Fig. 2, if \( P_i \) indicates the feature point of a planar object before the 3D movement, and \( Q_i \) indicates that after 3D movement, 3D motion can be expressed by (assuming the object is rotated by \( R \) first, and then it is translated by \( t \)):

\[
Q_i = R P_i + t
\]  

(3-64)

Here, \( i \) is the index of the feature points.

As a result, the square of the distance between \( Q_i \) and the optical center is

\[
Q_i Q_i' = (R P_i + t)' (R P_i + t) = P_i' R' R P_i + P_i' R' t + t' R P_i + t' t
\]

\[
= P_i' P_i + 2 t' (R P_i) + t' t
\]

(3-65)

The difference in distance between corresponding feature points is

\[
Q_i Q_i' - P_i' P_i = 2 t' (R P_i) + t' t
\]

\[
= 2 t' (R P_i + t) - t' t
\]

\[
= t' [2 (R P_i + t) - t]
\]

\[
= t' [2 Q_i - t]
\]  

(3-66)
Fig. 6 Possible perspective projection of one object with two kinds of 3D motion

(a) 3D motion 1

(b) 3D motion 2
It is necessary to guarantee Dependence 1 is one criterion of obtaining a unique solution by homography decomposition. Thus, at least one pair of corresponding feature points has a different sign in all the feature points in Eq. (3-66). Therefore, the constraint condition on distance in Dependence 1 has been changed into the constraint condition on angle \( \beta_i \) between vector \( t \) and vector \( 2Q_i - t \). If there is at least one \( \beta_i \) is larger than 90°, and the others are not, the unique solution will be obtained.

Fig. 6 illustrates possible perspective projection geometry of one object with two kinds of 3D motion. There is only one static camera \( O \) acting as the viewpoint. Two 3D feature planes, \( \Pi_1 \) and \( \Pi_2 \), are the same planar objects before and after 3D motion. \( P_i \) is an arbitrary 3D feature point on plane \( \Pi_1 \), and \( Q_i \) is the corresponding feature point on plane \( \Pi_2 \).

In Fig. 6 (a), it is obvious that motions such those as in Fig. 6 (a) cannot satisfy Dependence 1, because

\[
\begin{align*}
\beta_1 < 90^\circ &\Rightarrow Q_iQ - P_iP > 0 \Rightarrow |OP_1| < |OQ_1| \\
\beta_2 < 90^\circ &\Rightarrow Q_i'Q - P_i'P > 0 \Rightarrow |OP_2| < |OQ_2|.
\end{align*}
\]

However, if the object moves by another set of 3D motion parameters, it is possible to obtain the configuration in Fig. 6 (b). This situation makes it much more possible to satisfy Dependence 1, because

\[
\begin{align*}
\beta_1 < 90^\circ &\Rightarrow Q_iQ - P_iP > 0 \Rightarrow |OP_1| < |OQ_1| \\
\beta_2 > 90^\circ &\Rightarrow Q_i'Q - P_i'P < 0 \Rightarrow |OP_2| > |OQ_2|.
\end{align*}
\]

Thus, the solution ambiguity problem can be avoided in these kinds of situations.

Therefore, solution ambiguity to a size fixed object (a feature region) depends on 3D motion parameters. A unique solution can be conditionally obtained. The “Dependence on 3D Motion Parameters” for solution ambiguity has been proved geometrically.
3.2 Dependence on Object Size of Solution Ambiguity by Homography Decomposition

In this section, I will discuss another property of the solution ambiguity problem by homography decomposition, when the feature region size is not fixed. Meanwhile the geometry proof will also be available in this section.

Fig. 5 in Section 2.5 outlines one of various possible situations for these four estimated normal vectors under constraint \( P \cdot n > 0 \). As seen in Fig. 5, \( n_1 \) and \( n_2 \) are on a straight line and in opposite directions. So are \( n_3 \) and \( n_4 \). If the ambiguity problem in homography matrix decomposition occurs under constraint \( P \cdot n > 0 \), one of the possibly remaining normal vectors will be \( n_1 \) or \( n_2 \), and the other will be \( n_3 \) or \( n_4 \). Furthermore, the true estimated result must be one of these two remaining normal vectors.

From the general algebraic expression of homography matrix

\[
H = R + \frac{tn'}{d},
\]

it is obvious that \( H \) is only affected by 3D motion parameters (rotation matrix \( R \), translation vector \( t \), normal vector \( n \) of initial object plane \( \Pi_1 \), and distance \( d \) between object plane and camera), i.e.

\[
H = f(R, \ t, \ n, \ d).
\]

Therefore, size of the feature region will not affect the homography matrix at all. Once the 3D motion parameters are fixed, altering the size of the feature region will not change homography matrix \( H \).

Thus, if the ambiguous solutions are obtained in estimating 3D motion parameters by homography decomposition, the geometric angle between the two estimated normal vectors, which are the remaining ambiguous solutions, does not change. In a word, if the 3D motion parameters are fixed, the homography matrix will be fixed, and the angle between the ambiguous estimated normal vectors of the initial object plane \( \Pi_1 \) will be fixed, too.
Fig. 7 Geometric positions of initial feature plane and estimated normal vectors
Fig. 7 illustrates two kinds of possible geometric relations for \( \mathbf{n}_i \) and \( \mathbf{P}_i \) \((i=1,2)\). There is only one static camera \( \mathbf{O} \) acting as the viewpoint. The 3D feature plane, \( \Pi_1 \), is the position of the object before 3D motion. \( \mathbf{P}_i \) is an arbitrary 3D feature point on plane \( \Pi_1 \).

Because one of these two remaining estimated normal vectors should be the true solution, it is marked as \( \mathbf{n}_{true} \) in Fig. 7 (a) and (b); the other will be a false solution, so it is marked as \( \mathbf{n}_{false} \) in Fig. 7 (a) and (b). Meanwhile, the angles between \( \mathbf{n}_{true} \) and \( \mathbf{n}_{false} \) in Fig. 7 (a) and (b) are the same, because of the assumption that the 3D motion parameters in these two Figures are not changed.

Obtaining a unique solution means that \( \mathbf{n}_{false} \) can be eliminated by \( \mathbf{P}_i \cdot \mathbf{n}_{false} > 0 \) to all of the feature points; meanwhile, \( \mathbf{n}_{true} \) will remain by \( \mathbf{P}_i \cdot \mathbf{n}_{true} > 0 \) to all of the feature points. That means, at least one of inner products between \( \mathbf{n}_{false} \) and \( \mathbf{P}_i \) should be negative which has a different sign from the others, and all of inner products between \( \mathbf{n}_{true} \) and \( \mathbf{P}_i \) should be positive.

From Fig. 7 (a), it is obvious that the solution ambiguity problem will occur, because

\[
\begin{align*}
\text{Angle}(\mathbf{n}_{true}, \mathbf{P}_1) < 90^\circ & \Rightarrow \mathbf{n}_{true}^t \mathbf{P}_1 > 0 \quad \Rightarrow \mathbf{n}_{true} \\
\text{Angle}(\mathbf{n}_{true}, \mathbf{P}_2) < 90^\circ & \Rightarrow \mathbf{n}_{true}^t \mathbf{P}_1 > 0 \quad \Rightarrow \mathbf{n}_{true} \\
\text{Angle}(\mathbf{n}_{false}, \mathbf{P}_1) < 90^\circ & \Rightarrow \mathbf{n}_{false}^t \mathbf{P}_1 > 0 \quad \Rightarrow \mathbf{n}_{false} \\
\text{Angle}(\mathbf{n}_{false}, \mathbf{P}_2) < 90^\circ & \Rightarrow \mathbf{n}_{false}^t \mathbf{P}_1 > 0 \quad \Rightarrow \mathbf{n}_{false}
\end{align*}
\]

(3-71)

Both of \( \mathbf{n}_{true} \) and \( \mathbf{n}_{false} \) will be obtained as the real estimated normal vector by homography decomposition.

However, under the same 3D motion parameters, if the feature region is enlarged as in Fig. 7 (b), the unique solution can be obtained, because

\[
\begin{align*}
\text{Angle}(\mathbf{n}_{true}, \mathbf{P}_1) < 90^\circ & \Rightarrow \mathbf{n}_{true}^t \mathbf{P}_1 > 0 \quad \Rightarrow \mathbf{n}_{true} \\
\text{Angle}(\mathbf{n}_{true}, \mathbf{P}_2) < 90^\circ & \Rightarrow \mathbf{n}_{true}^t \mathbf{P}_1 > 0 \quad \Rightarrow \mathbf{n}_{true} \\
\text{Angle}(\mathbf{n}_{false}, \mathbf{P}_1) > 90^\circ & \Rightarrow \mathbf{n}_{false}^t \mathbf{P}_1 < 0 \Rightarrow \times \\
\text{Angle}(\mathbf{n}_{false}, \mathbf{P}_2) < 90^\circ & \Rightarrow \mathbf{n}_{false}^t \mathbf{P}_1 > 0 \quad \Rightarrow \times
\end{align*}
\]

(3-72)
\( n_{\text{false}} \) will be eliminated since the angle between \( n_{\text{false}} \) and \( P_1 \) is larger than 90°, while that between \( n_{\text{false}} \) and \( P_2 \) is smaller than 90°. Finally, only \( n_{\text{true}} \) will be obtained as the real estimated normal vector by homography decomposition.

This property of the solution ambiguity problem is described as Dependence 2:

**Dependence 2**

“\( \text{To a set of fixed 3D motion parameters, if the size of the feature region (object) is large enough, or the range of the 3D movement is relatively small compared to the size of the feature region (object), the solution ambiguity problem obtained by homography decomposition will be avoided.} \)”

Actually, Dependence 2 can also be proved under the theory of Dependence 1 which is proposed in last section.

To obtain the unique solution, it is necessary to guarantee Dependence 1 is satisfied. Thus, at least one pair of corresponding feature points has a different sign in all the feature points in Eq. (3-66). Therefore, the constraint condition on distance in Dependence 1 has been changed into the constraint condition on angle \( \beta_i \) between vector \( t \) and vector \( 2Q_i - t \). If there is at least one \( \beta_i \) is larger than 90°, and the others are not, the unique solution will be obtained.

Fig. 8 illustrates possible perspective projection geometry of one set of 3D motion parameters with two kinds of feature region sizes. There is only one static camera \( O \) acting as the viewpoint. Two 3D feature planes, \( \Pi_1 \) and \( \Pi_2 \), are the same planar objects before and after 3D motion. \( P_i \) is an arbitrary 3D feature point on plane \( \Pi_1 \), and \( Q_i \) is the corresponding feature point on plane \( \Pi_2 \).

It is obvious that motions such those as in Fig. 8 (a) cannot satisfy Dependence 1, because

\[
\begin{align*}
\beta_1 < 90^\circ \Rightarrow Q'_iQ - P'_iP > 0 \Rightarrow |OP_1| < |OQ_1| \\
\beta_2 < 90^\circ \Rightarrow Q'_2Q - P'_2P > 0 \Rightarrow |OP_2| < |OQ_2|
\end{align*}
\]

(3-67)
Fig. 8 Possible perspective projection of an object (the same position, but different size)
Therefore, the ambiguous solutions will be obtained in the situation as shown in Fig. 8 (a).

However, if the feature region is enlarged and the 3D motion parameters are kept unchanged, the geometrical relations among the feature points in Fig. 8(a) will be changed, as shown in Fig. 8 (b). This situation makes it much more possible to satisfy Dependence 1, because

\[
\begin{align*}
\beta_1 < 90^\circ & \Rightarrow Q_1Q - P_1P > 0 \Rightarrow |OP_1| < |OQ_1| \\
\beta_2 > 90^\circ & \Rightarrow Q_2Q - P_2P < 0 \Rightarrow |OP_2| > |OQ_2|
\end{align*}
\]  (3-68)

Therefore, the unique solution will be obtained in the situation as shown in Fig. 8 (6).

Based on the theory of Dependence 1, Dependence 2 has also been proved geometrically. The latter geometrical proof of Dependence 2 does not only verification the reliability of Dependence 2, but also reveals the mutual relationship.

Therefore, based on what has been discussed above, to a set of fixed 3D motion parameters, solution ambiguity by homography decomposition depends on the object size. A unique solution can still be conditionally obtained. This property is defined as “Dependence on Object Size” of solution ambiguity.

In this chapter, I do not only prove that the solution ambiguity problem in estimating the 3D motion parameters by homography decomposition is not inevitable. It depends on two kinds of factors: (1) when the object size is fixed, it depends on changing the 3D motion parameters; (2) when the 3D motion parameters are fixed, it depends on changing the size of the object. I define them as “Dependence on 3D Motion Parameters” and “Dependence on Object Size” of solution ambiguity respectively. Meanwhile, the strict theoretical and geometrical proofs are operated in this chapter, too.
Chapter 4  Constraint Condition for Homography Decomposition

In last chapter, I proposed “Dependence on 3D Motion Parameters” and “Dependence on Object Size” of solution ambiguity in estimating 3D motion parameters by homography decomposition. However, the two dependences just reveal the fact that the solution ambiguity problem is not inevitable, and they show a trend in which the ambiguous solution in estimating 3D motion parameters by homography decomposition will be more possible eliminated. Thus, it does not mean that the unique solution can be obtained definitely with only two pieces of image to estimate the 3D motion parameters by homography decomposition.

As a result, the extra constraint condition is still necessary in the applications. Even though there has been quite a few constraint conditions proposed until now, such as the information among the time series frames, the information from the extra viewpoint, and so on, I will proposed another kinds of constraint condition which is still based on only two pieces of images.

Most of the previous researches were focus on only one plane. As a result, the solution ambiguity problem occurs easily by only two pieces of image. However, considering that most of the approximate shapes of the objects are easily to be known (such as a cube, a tetrahedron and so on), the angle between two planes must be a kind of strong constraint for the shape reconstruction or the 3D motion parameters estimation. Therefore, I propose the “Two Planes Constraint” to resolve the solution ambiguity problem in estimating 3D motion parameters by homography decomposition with only two pieces of images.

Fig. 9 illustrates the geometrical relationship of the 3D motion by two linked planes. There is only one static camera 0 acting as the viewpoint.  and  are the two linked planes, and the angle between these two planes,  and , is .  is projected on the image plane as , and  is projected on the image plane as . In my research,  and  are rotated and translated in 3D space respectively by the
same 3D motion parameters. After the 3D movement, \( \Pi_{11} \) moved to the poison of \( \Pi_{21} \), \( \Pi_{12} \) moved to the poison of \( \Pi_{22} \). Note that the angle between the two linked planes will not be changed during the 3D movement. Thus, the angle between these two planes, \( \Pi_{21} \) and \( \Pi_{22} \), is still \( \theta \). Meanwhile, \( \Pi_{21} \) is projected on the image plane as \( \pi_{21} \), and \( \Pi_{22} \) is projected on the image plane as \( \pi_{22} \).

Fig. 9 Geometry of 3D motion by two linked planes

In my research, one homography matrix, \( H_1 \), will be obtained by the corresponding feature points on \( \pi_{11} \) and \( \pi_{21} \), and the other homography matrix, \( H_2 \), will be obtained by the corresponding feature points on \( \pi_{12} \) and \( \pi_{22} \). Totally, two homography matrices will be calculated in my research.

Then, these two homography matrices will be decomposed respectively. If the solution ambiguity problem occurs in both decompositions of the two homography matrices, \( H_1 \) and \( H_2 \), totally four estimated normal vectors of the initial object plane, \( \Pi_{11} \) and \( \Pi_{12} \). The two sets of the ambiguous estimated normal vectors for the two
linked planes can combined into four possible situations. Comparing the angle between the two estimated normal vectors, $\mathbf{n}_i$ and $\mathbf{n}_j$ ($i=1,2$, $j=1,2$, $\mathbf{n}_i$ stands for the two estimated normal vectors of \( \Pi_{11} \), and $\mathbf{n}_j$ stands for the two estimated normal vectors of \( \Pi_{12} \)), and the angle $\theta$ between these two planes, the angle between the two real estimated normal vectors will be the supplementary angle of $\theta$, according to the invariance of the angle between two linked planes.

![Diagram illustrating the combination of ambiguous solutions from linked two planes.](image)

**Fig. 10 Combination of ambiguous solutions from linked two planes**

Fig. 10 illustrates the geometrical relation of the solution ambiguity problem occurs in both decompositions of the two homography matrices, $\mathbf{H}_1$ and $\mathbf{H}_2$, for the two linked planes. ($\mathbf{n}_{true1}$ and $\mathbf{n}_{false1}$ are the ambiguous solutions of $\Pi_{11}$, and $\mathbf{n}_{true2}$ and $\mathbf{n}_{false2}$ are the ambiguous solutions of $\Pi_{12}$. Actually, it is not clear which is the real estimated normal vector or which is the fake estimated one for either of two planes.) By combining these two set of ambiguous solutions, four kinds of combinations of the estimated normal vectors will be obtained finally:

$$\{\mathbf{n}_{true1}, \mathbf{n}_{false1}\} \times \{\mathbf{n}_{true2}, \mathbf{n}_{false2}\}$$

Then the angle $\phi_i$ ($i=1,2,3,4$) between the two estimated normal vectors of any
The combination which can make Eq. (4-2) to obtain the minimum evaluation value \( \eta \), is considered the combination of both two real estimated normal vectors for the two linked planes, \( \Pi_{11} \) and \( \Pi_{12} \).

The constraint condition of two linked planes performs very strong and effective filtration to the outliers during the simulation experiment which is based on a quite huge database. Even though the combination which contains the false estimated normal vectors can obtain the minimum valuation value \( \eta \) occasionally, this probability is smaller than 0.3\%. It is too rare to conceal the fact that the information of two linked planes is a kind of effective constraint condition to the solution ambiguity problem in estimating 3D motion parameters by homography decomposition with only two pieces of image.
Chapter 5   Simulation Experiment and Results

In my research, lots of work has been done to reveal the truth on the linear 3D motion parameters estimation method, against to the misunderstanding for a long history. Therefore, the main purpose of my research is not to propose a kind of new approach to estimate 3D motion parameters more accurately by lower cost. In my research, I mainly intent to propose a more complete description of the existing homography decomposition method in estimating 3D motion parameters by only two pieces of images.

Therefore, the experiments are set to test whether my theories are correct, but not to test how the proposed approach performs, such as the accuracy, the speed, the robustness to the noisy and so on. As a result, all testaments are operated by the simulation experiments. Although the noise affect has been ignored, and the real data are lack to be tested, the results from the simulation experiments which are based on the huge database are sufficient to prove correctness of my theories.

In this chapter, I will discuss the simulation experiments by three themes: the convergence of the iterative rotation center estimating method; the two dependences of solution ambiguity problem in estimating 3D motion parameters by homography decomposition method; the constraint of two linked planes.

5.1 The Convergence of the Iterative Rotation Center Estimating Method

To test the convergence of the iterative rotation center estimating method which is proposed in Section 2.4, the experiments is set to simulate the situation which a planar object is rotated on an arbitrary spot in the 3D space. In this way, all kinds of possible situations of the initial object plane \( \Pi_1 \) can be simulated.
Fig. 11 illustrates the geometrical relationship of this simulation experiments. There is only one static camera O acting as the viewpoint. To decide the object position in the 3D space, I first set the angle $\theta$ between Z axis and Z’ axis manually. All of the candidates of angle $\theta$ are $\{-30^\circ, -20^\circ, -10^\circ, 0^\circ, 10^\circ, 20^\circ, 30^\circ\}$. Then I point an arbitrary poison on Z’ axis in the space as the rotation center $O_2$ of the object, i.e. the geometry center of the feature center. Then I set the 2D coordinate of the four feature points to decide the shape and the size of the feature region. The size and the shape of the feature region will not be changed during the whole simulation experiment. In my research I set them as (-200, -160), (200, -160), (-140, 160), and (140, 160). The unit is millimeter. Based on the rotation center $O_2$ which I have already set, the space position of $\Pi_1$ can be determined, as the black solid line in Fig. 11. Then $\Pi_1$ is projected onto the image plane, and the image $\pi_1$ can be obtained, as
the green solid line in Fig. 11. $\Omega_1$ and $\Omega_2$ are the virtual planes which were mentioned in Section 2.4. Then it is possible to estimate the coordinate of real rotation center $O_1$ by the iterative method which are proposed in Section 2.4.

After one time estimating simulation is successful, $\Pi_1$ will be rotated to another position $\Pi'_1$, as the black dotted line in Fig. 11. By the new $\Pi'_1$, the new estimation steps above will be repeated. On each candidate position of the $Z'$ axis, the estimating simulation will be repeated ten times. Then the average of value for each simulation will be the final result for that position. Finally, the results are shown in Fig. 12.

![Distance Convergence for Iteration](image)

**Fig. 12 Distance Convergence for Iteration**

In Fig. 12, the horizontal axis represents the index of the iteration; the vertical axis represents the distance between the rotation center on the image plane $o_j$ and the average coordinate $o_j'$ of the feature points in the 3D space. The parameter shows the angle $\theta$ between Z axis and $Z'$ axis. Each curve in Fig. 12 shows that the distance converges to zero rapidly even when the initial distance is big. Eventually, all the iteration for different angle can be finished in 5 iterations. The results show that my rotation center estimating method is really convergent and low cost, even though it is a kind of iterative method. As a result, it is clear that the computing cost will not be affect very much to estimate the real rotation center and the real translation vector $t$. 

75
during estimation of the 3D motion parameter by homography decomposition, by the iterative rotation center estimating method proposed in Section 2.4.

5.2 The Two Dependences of Solution Ambiguity Problem in Estimating 3D Motion Parameters by Homography Decomposition Method

In this section, the experiments will be operated to simulate the all of the possible 3D movements in the space. For each simulation the 3D motion parameters will be estimated by homography decomposition method. Thus, hundred thousands of simulation data constructed a very huge database for my testing experiments. Therefore, it is possible to believe that the experiments results are reliable enough.

I will introduce the setting for my simulation experiment. Fig. 12 illustrates the
geometric relations in my simulation experiments. First, I assume an initial position of feature plane $\Pi_0$, which is on the Image Plane, and the optical axis passes through the rotation center (geometry center) of $\Pi_0$. It is then transformed by one kind of 3D movement ($R_1$, $t_1$) to position $\Pi_1$, where observed image $\pi_1$ can be obtained by projection. $\Pi_0$ is simultaneously transformed by another kind of 3D movement ($R_2$, $t_2$) to position $\Pi_2$, where observed image $\pi_2$ can also be obtained. $P_i$ is the arbitrary 3D feature point on the plane $\Pi_i$, and $Q_i$ is the corresponding feature point on plane $\Pi_2$. $p_i$ and $q_i$ are the two corresponding feature vectors on the Image Plane, $\pi_1$ and $\pi_2$, respectively.

The purpose of the simulation experiments was to estimate the 3D motion parameters from $\Pi_1$ to $\Pi_2$ ($R_{12}$, $t_{12}$) using information only from two images $\pi_1$ and $\pi_2$ ($p_i$ and $q_i$) by the homography matrix decomposition algorithm. According to the estimated results, estimated rotation matrix $\hat{R}_{12}$ is exactly the same as rotation matrix $R_{12}$ between $\Pi_1$ and $\Pi_2$. However, because depth information on the object is unknown, estimated translation vector $\hat{t}_{12}$ is proportional to translation vector $t_{12}$ from $\Pi_1$ to $\Pi_2$.

![Fig. 14 A simulation image for the 3D movement](image)

Fig. 14 illustrates the situation of the example of the simulated image, in which I show the feature region before and after the 3D movement in one piece of image for convenience. The quadrangle “$p_1 p_2 p_3 p_4$” of blue solid line represents the feature
region which consists of the feature points projected from the object plane before the 3D movement; the quadrangle “q₁ q₂ q₃ q₄” of red dotted line represents the feature region which consists of the feature points projected from the object plane after the 3D movement. The black dotted lines show the horizontal and vertical middle lines of the image in order to express the image center. The simulation experiment is to decompose the homography matrix, which is obtained by pᵢ and qᵢ (i = 1, 2, 3, 4), in order to estimated the 3D motion parameter in the space.

Two kinds of shapes for the feature region were used in my experiments: the first was a symmetrical quadrangle, and the second was an asymmetrical quadrangle. Because at least four pairs of corresponding feature points are needed to obtain the homography matrix between two images, four corner points of the feature region are be used in the simulation experiments. The simulation data setting must guarantee that the object is transformed within a range in which the front surface of the object is always visible before and after 3D transformation. Thus, the four feature points on Π₀ are as follows:

<table>
<thead>
<tr>
<th>Feature point coordinates on Π₀ of symmetrical quadrangle</th>
<th>Feature points coordinates on Π₀ of asymmetrical quadrangle</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_{10} : (-200, -160) )</td>
<td>( P_{10} : (-200, -140) )</td>
</tr>
<tr>
<td>( P_{20} : (200, -160) )</td>
<td>( P_{20} : (250, -100) )</td>
</tr>
<tr>
<td>( P_{30} : (-140, 160) )</td>
<td>( P_{30} : (-120, 2000) )</td>
</tr>
<tr>
<td>( P_{40} : (140, 160) )</td>
<td>( P_{40} : (180, 120) )</td>
</tr>
</tbody>
</table>

Unit: pixel

Focal length \( f \) is 400 pixels. The image is assumed to be 640 × 480 pixels.

All 12 parameters for \( R_1, t_1, R_2, \) and \( t_2 \) (three 3D rotation angles around X, Y,
and Z axes, respectively, \( \alpha, \beta, \gamma \), for each rotation matrix, \( R_1 \) or \( R_2 \); three 3D translation components along X, Y, and Z axes, respectively, \( \Delta X, \Delta Y, \Delta Z \), for each translation vector, \( t_1 \) or \( t_2 \).) were changed to simulate the 3D movements between \( \Pi_1 \) and \( \Pi_2 \). The parameters are setting among following range, as shown in Table 7.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Changing Range</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>3D Rotation (( R_1 ) and ( R_2 ))</strong> (Unit: °)</td>
<td></td>
</tr>
<tr>
<td>( \alpha ) (around X axis)</td>
<td>-45 ~ 45</td>
</tr>
<tr>
<td>( \beta ) (around Y axis)</td>
<td>-45 ~ 45</td>
</tr>
<tr>
<td>( \nu ) (around Z axis)</td>
<td>-45 ~ 45</td>
</tr>
<tr>
<td><strong>3D Translation (( t_1 ) and ( t_2 ))</strong> (Unit: pixel)</td>
<td></td>
</tr>
<tr>
<td>( \Delta X )</td>
<td>-80 ~ 80</td>
</tr>
<tr>
<td>( \Delta Y )</td>
<td>-80 ~ 80</td>
</tr>
<tr>
<td>( \Delta Z )</td>
<td>-80 ~ 80</td>
</tr>
</tbody>
</table>

In total, nearly \( 6 \times 10^6 \) sets of simulation data could be obtained for each shape of the feature region.

First, the symmetry quadrangle feature region was used. The percentage of one solution cases and two solution cases is outlined in Fig. 15 obtained by using the linear homography decomposition algorithm proposed by Zhang and Hanson and the general constraints in which all feature points should satisfy \( P \cdot n > 0 \).

The conclusion stated by Zhang and Hanson is widely accepted. Namely that, “…this ambiguity of the two solutions cannot be resolved if there is no further a priori information available…” [38]. However, the results from my simulation experiments revealed that the one solution cases have a high percentage of occurrences without any a priori information.

The asymmetrical feature region was also tested in my research. The ratio of cases without the solution ambiguity problem was exactly the same as that for the
symmetry quadrangle feature region. Therefore, the simulation experiments result shown that the solution ambiguity problem is not inevitable, even though only two pieces of images and the solution reality constraint can be used.

![Occurrence ratio of one solution cases and two solution cases](image)

Fig. 15 Occurrence ratio of one solution cases and two solution cases

I then investigated the dependence of the solution ambiguity problem on 3D motion parameters (Dependence 1), which was mentioned in Section 3.1. I define parameter \( d / d' \) to indicate how many feature points satisfy Dependence 1,

\[
d = \frac{1}{n} \sum_{i=1}^{n} P'_i P_i
\]

(5-1)

and

\[
d' = \frac{1}{n} \sum_{i=1}^{n} Q'_i Q_i
\]

(5-2)

If there are more feature points satisfying Dependence 1, the value of \( d / d' \) will be closer to 1. Otherwise, the value of \( d / d' \) will be farther from 1.

Fig. 16 plots the results obtained from the simulation experiments. The horizontal axis is \( d / d' \) and the right vertical axis is the number of ambiguous solutions or unique solutions. The left vertical axis is the ratio of ambiguous (or unique) solutions in the total estimation results. The green bars indicate the number of
unique solutions to the different $d/d'$, and the blue bars indicate the number of ambiguous solutions to different $d/d'$. The red curve with closed triangles (▲) indicates the ratio of unique solutions to different $d/d'$ and the purple curve with closed squares (■) indicates the ratio of ambiguous solutions to different $d/d'$.

![Fig. 16 Change in numbers of solutions due to dependence of solution ambiguity on 3D motion parameters](image)

From Fig. 16, the number of unique solutions reaches maximum in the range between 0.9 and 1, and almost all the estimates obtain a unique solution. Otherwise, the ratio of unique solution will decrease, and the ratio of the ambiguous solutions will increase. Therefore, the simulation results reveal that if Dependence 1 is satisfied, a unique solution tends to be obtained by homography decomposition. The correctness of “Dependence on 3D Motion Parameters” is given with the simulation results.

I next investigated the dependence on object size (Dependence 2) of the solution ambiguity problem explained in Section 3.2. I changed the size of feature region $\Pi_0$ by a scale factor, using the same 3D motion parameter settings. Scale factor is the parameter by which the size of the planar object in the space can be changed. The scale factor was not only valued from 2 to 0.25 by 0.25, but also set at 0.1 and 0.01.
The change in the ratios of one solution and two solution cases are given in Fig. 17, where the horizontal axis is the scale factor for the feature region, and the vertical axis is the ratio of ambiguous (or unique) solutions in the total estimated results. The curve with closed circles (●) plots the ratio of unique solutions to the various scale factors, and the curve with closed triangles (▲) plots the ratio of ambiguous solutions to various scale factors.

![Fig. 17 Change in numbers of solutions due to dependence of solution ambiguity on object size](image)

As shown in Fig. 17, when the scale factor is equal to 0.01, there are almost nine times more two solutions than one solution. As the feature region size is increased, the curve for the one solution cases rises, and the curve for the two solutions cases descends. Thus, if the feature region size is increased, situations without ambiguity will increase. In contrast, ambiguity situations will decrease. Therefore, the simulation results demonstrated the accuracy of “Dependence on Object Size” of solution ambiguity which are proposed in Section 3.2.

Therefore, the huge number of simulation results not only points out that solution ambiguity is not inevitable, but it also coincides with the two dependences of solution ambiguity I proposed in this paper: the “Dependence on 3D Motion Parameters”
and the “Dependence on Object Size” of solution ambiguity problem in estimating 3D motion parameters by homography decomposition method.

5.3 Constraint Condition of Joined Two Planes

Based on the constraint condition which is proposed in Chapter 4, I set the experiments to simulate the 3D movement of the two linked planes.

![Diagram](image.png)

**Fig. 18** Geometric relations in simulation for two linked object planes

Fig. 18 illustrates the geometric relations in my simulation experiments. First, I assume an initial position of feature plane $\Pi_{01}$, which is on the Image Plane, and the optical axis passes through the rotation center (geometry center) of $\Pi_{01}$. Then on one side of $\Pi_{01}$, the initial position of another plane $\Pi_{02}$ is set, which is linking to $\Pi_{01}$. By rotating $\Pi_{02}$ it is possible to determine the angle $\theta$ between $\Pi_{01}$ and $\Pi_{02}$. And the angle $\theta$ will not be changed during the whole simulation experiments. $\Pi_{01}$ and $\Pi_{02}$ are then transformed by one kind of 3D movement ($R_x, t_x$) to positions $\Pi_{11}$ and $\Pi_{12}$.
respectively, where observed images $\pi_{11}$ and $\pi_{12}$ can be obtained by projection. $\Pi_{01}$ and $\Pi_{02}$ are simultaneously transformed by another kind of 3D movement ($R_2$, $t_2$) to positions $\Pi_{21}$ and $\Pi_{22}$ respectively, where observed images $\pi_{21}$ and $\pi_{22}$ can be obtained by projection. One homography matrix, $H_1$, will be obtained by the corresponding feature points on $\pi_{11}$ and $\pi_{21}$, and the other homography matrix, $H_2$, will be obtained by the corresponding feature points on $\pi_{12}$ and $\pi_{22}$. Totally, two homography matrices will be calculated in my research.

The purpose of the simulation experiments is to estimate the 3D motion parameters from $\{\Pi_{11}, \Pi_{12}\}$ to $\{\Pi_{21}, \Pi_{22}\}$ respectively ($R_x$, $t_x$) using information from $\{\pi_{11}, \pi_{12}\}$ and $\{\pi_{21}, \pi_{22}\}$ by the homography matrix decomposition algorithm. If the solution ambiguity problem occurs, the constraint condition of two linked planes will be used to obtain the unique estimated solution.

As the same as the feature point setting in last section, the four corner points of the feature region are used. The simulation data setting must guarantee that the object is transformed within a range in which the front surface of the object is always visible before and after 3D transformation. Thus, the four feature point coordinates on $\Pi_{01}$ are: (-200, -160), (200, -160), (-140, 160), and (140, 160). The four feature point coordinates on $\Pi_{02}$ are set by rotation and translation result of $\Pi_{01}$. The units of the coordinates are pixels. Focal length $f$ is 400 pixels. The image is assumed to be 640 $\times$ 480 pixels.

In the simulation experiments, four kinds of the angle $\theta$ between $\Pi_{01}$ and $\Pi_{02}$ are tested: $150^\circ$ (the angle between the normal vectors of $\Pi_{01}$ and $\Pi_{02}$ is $30^\circ$ ); $135^\circ$ (the angle between the normal vectors of $\Pi_{01}$ and $\Pi_{02}$ is $45^\circ$ ); $120^\circ$ (the angle between the normal vectors of $\Pi_{01}$ and $\Pi_{02}$ is $60^\circ$ ); $90^\circ$ (the angle between the normal vectors of $\Pi_{01}$ and $\Pi_{02}$ is $90^\circ$ ). Finally, because the exact same results are obtained, only the situation which the angle $\theta$ between $\Pi_{01}$ and $\Pi_{02}$ is $135^\circ$ will be discussed in this paper, as a typical example.

All 12 parameters for $R_1$, $t_1$, $R_2$, and $t_2$ (three 3D rotation angles around X, Y, and Z axes, respectively, for each rotation matrix, $R_1$ or $R_2$; three 3D translation components along X, Y, and Z axes, respectively, for each translation vector, $t_1$ or $t_2$)
were changed to simulate the 3D movements between $\Pi_{1i}$ and $\Pi_{2i}$ ($i = 1, 2$). The rotation angles changed from $-45^\circ$ to $45^\circ$, and the translation parameters changed from -100 pixels to 100 pixels, to guarantee the projected feature region would remain inside the image. In total, nearly $6 \times 10^6$ sets of simulation data could be obtained for each shape of the feature region. However, the situations, which the same side of the object plane can not be observed before and after the 3D motion, are ignored.

Because in my research, the combination which can make Eq. (4-2) to obtain the minimum evaluation value $\eta$, is considered the combination of both two real estimated normal vectors for the two linked planes, $\Pi_{11}$ and $\Pi_{12}$, the unique solution is definitely obtained. However, the combination which contains the false estimated normal vectors can obtain the minimum valuation value $\eta$ occasionally, the ratio of the combinations is going to be tested, which are both real estimated solutions, among the all data. Meanwhile, the size of feature region $\Pi_{0i}$ ($i = 1, 2$) are changed by the scale factor $k$, using the same 3D motion parameter settings. The scale factor $k$ was not only valued from 2 to 0.25 by 0.25, but also set at 0.1.

The change in the ratios of real estimated solution combinations for various object sizes are given in Fig. 19, where the horizontal axis is the scale factor for the feature region size.
feature region, and the vertical axis is the ratio of solutions in the total estimated results. The dotted curve with closed diamond (♦) plots the ratio of unique solutions without the two plane constraint to the various scale factors, and the dotted curve with closed squares (■) plots the ratio of ambiguous solutions without the two plane constraint to various scale factors; the solid curve with closed triangles (▲) plots the ratio of ambiguous solutions to various scale factors, which constrained by the two linked planes condition.

From Fig. 19, it is clear that, if only one plane is available to estimate the 3D motion parameters by homography decomposition method, the solutions ambiguity problem shows the same dependence as Fig. 17: if the feature region size is increased, situations without ambiguity will increase; in contrast, ambiguity situations will decrease.

If the constraint condition of the two linked planes is used, ratio of the unique real solution will be invariant to the object size, and it is extremely near to the 100%. Meanwhile, the ratio of the outliers is too rare to conceal the effectiveness of the constraint which is proposed in my research.

In summary, the theories, the iterative method and the constraint condition which are proposed in this paper have been completely proved by the simulation experiments.
Chapter 6  Conclusion

In my research, I accomplished a deeper understanding of the solution ambiguity problem in the estimation of 3D motion parameters obtained by homography decomposition. I proposed two kinds of dependencies for solution ambiguity obtained by homography decomposition that have been ignored or misunderstood for quite a long time. What are proposed in this paper indicates the solution ambiguity problem is not inevitable and a unique solution can be conditionally obtained, even though only two pieces of images are available. Meanwhile, a kind of constrain condition is also proposed to guarantee the real unique estimated solution can be obtained in the applications.

First, the paper pointed out the dependence of solution ambiguity on 3D motion parameters; if all distances between the feature points and the camera do not become closer after 3D motion, the solution ambiguity problem can be avoided. This conclusion has been geometrically and theoretically derived and this is the first time this dependence has been mentioned. However, because the constraint for the theoretical proof is more common than that in the previous work, this theory should be more reliable.

Second, this paper discussed the dependence of solution ambiguity on object size and explained it with geometry; if the feature region is large enough, or the range of 3D movements is relatively small compared to the size of the feature region, the solution ambiguity problem can be avoided by using homography decomposition. This is the first time this dependence has been mentioned until now.

Third, this paper propose a kind of constraint condition of two linked planes, in order to guarantee the unique real estimated solution can be obtained in the applications. Meanwhile, the effectiveness of the constraint condition has been tested by the simulation experiments. The ratio of the outliers is smaller than 0.3‰.

This paper clarified the mechanism for solution ambiguity obtained by homography decomposition in the estimation of 3D motion parameters. Both
dependencies on 3D motion parameters and object size have been theoretically and experimentally verified. Therefore, all the results which are obtained in my simulation experiments should have been sufficiently accurate so that a demonstration with real images was not necessary in this paper.

In this paper, I neither proposed a new approach to solving the solution ambiguity problem obtained by homography decomposition, nor proposed a new linear method without the solution ambiguity problem to estimate 3D motion parameters. I only pointed out two kinds of dependencies, based on a deeper understanding of the solution ambiguity problem obtained by homography decomposition.

I would like to do further research in future work on application of homography decomposition method based on my theories, and I also intend to research on the relation between the homography decomposition method and other linear estimation method, such as the factorization method, in order to find the new linear estimation method for 3D motion parameters.
References


Publication List

Journal Paper


International Conferences


National Conferences


Conference of Hokuriku Chapters of Electrical Societies, Kanazawa, Japan, September, 2009.


