January 1989 UILU-ENG-89-2208

COORDINATED SCIENCE LABORATORY
College of Engineering

MOTION ANALYSIS I: BASIC THEOREMS, CONSTRAINTS, EQUATIONS, PRINCIPLES AND ALGORITHMS

Xiaoping Hu
Narendra Ahuja

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION: Unclassified

1b. RESTRICTIVE MARKINGS: None

2a. SECURITY CLASSIFICATION AUTHORITY

2b. DECLASSIFICATION / DOWNGRADING SCHEDULE

3. DISTRIBUTION / AVAILABILITY OF REPORT: Approved for public release; distribution unlimited

4. PERFORMING ORGANIZATION REPORT NUMBER(S): UILU-ENG-89-2208

5. MONITORING ORGANIZATION REPORT NUMBER(S)

6a. NAME OF PERFORMING ORGANIZATION: Coordinated Science Lab, University of Illinois

6b. OFFICE SYMBOL (If applicable): N/A

7a. NAME OF MONITORING ORGANIZATION

6c. ADDRESS (City, State, and ZIP Code): 1101 W. Springfield Ave. Urbana, IL 61801

7b. ADDRESS (City, State, and ZIP Code)

8a. NAME OF FUNDING / SPONSORING ORGANIZATION

8b. OFFICE SYMBOL (If applicable)

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER

8c. ADDRESS (City, State, and ZIP Code)

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification): Motion Analysis I: Basic Theorems, Constraints, Equations, Principles and Algorithms

12. PERSONAL AUTHOR(S): Hu, Xiaoping and Ahuja, Narendra

13a. TYPE OF REPORT: Technical

13b. TIME COVERED: FROM ________ TO ______

14. DATE OF REPORT (Year, Month, Day): January 1989

15. PAGE COUNT: 86

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number): motion, stereo vision, depth, surface, rigidity, planar motion solution

19. ABSTRACT (In this report we present some theorems and algorithms that might be basic for the motion and structure solution from plane and surface point correspondences. These theorems give the simplest linear solutions of motion so far. Almost every theorem presented here has been verified directly or indirectly by experiments. The mathematical proofs of the theorems are also very simple. We also review some other algorithms and point out the disadvantages in them. We list other constraints for a rigid motion and then point out that 4-point correspondences on a general surface in two views might determine the motion to a finite number of solutions. We argue that the traditional 8-point linear algorithms will not always promise a consistent solution and hence the robustness of them is limited. We propose a new criterion for optimal solution judgment and discuss the relationship between long-range motion and short-range motion. Our algorithms differ from the traditional linear algorithms in methodology but not only in techniques. Our solution is linear and globally optimal satisfying all constraints because we search in a reasonable space. Simulation results are presented. These results show that it is possible to achieve acceptable accuracy with the current camera resolution by our methods.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT: ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION: Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL

22b. TELEPHONE (Include Area Code)

22c. OFFICE SYMBOL

DD Form 1473, JUN 86 — Previous editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

Motion Analysis I:

Basic Theorems, Constraints, Equations, Principles and Algorithms

Xiaoping Hu

Narendra Ahuja

Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
1101 W. Springfield Ave.
Urbana, IL 61801, USA

email: xh, ahuja%uirvld.csl.uiuc.edu@uxc.cso.uiuc.edu

January, 1989

Abstract

In this paper we present some theorems and algorithms which might be basic for the motion and structure solution from plane and surface point correspondences. These theorems give the simplest linear solutions of motion so far. Almost every theorem presented here has been verified directly or indirectly by experiments. The mathematical proofs of the theorems are also very simple. We also review some other algorithms and point out the disadvantages in them. We list other constraints for a rigid motion and then point out that 4 point correspondences on a general surface in two views might determine the motion to a finite number of solutions. We argue that the traditional 8-point linear algorithms will not always promise a consistent solution and hence the robustness of them is limited. We propose a new criterion for optimal solution judgement and discuss the relatioship between long range motion and short range motion. Our algorithms differ from the traditional linear algotithms in methodology but not only in techniques. Our solution is linear and globally optimal satisfying all constraints because we search in a reasonable space. Simulation results are presented. These results show that it is possible to achieve acceptable accuracy with the current camera resolution by our methods.

Index: motion, stereo vision, depth, surface, rigidity, planar motion solution

1. Introduction

The motion problem is long unsolved because of the limited resolution of current cameras and the difficulty in feature matching. So far no algorithm shows that it works for real images taken by the real cameras. Because of the quantization error, most algorithms do not seem robust in motion parameter estimation. Hence many reseachers turn to use more than enough correspondences to get an optimal estimation in the sense of least square (Weng, 87). However we would like to point out in this paper that the current least square algorithms do not promise a robust solution. One problem is that the least square solution cannot guarantee that the more correspondences are used, the more accurate the the motion estimation will be. Aother problem is in that all these algorithms are two step linear algorithms which first solve an intermediate matrix or vector, say $\mathbf{M}$, and then solve the motion parameters **R** and **T** from **M**. It is well known that the motion problem is intrinsically a nonlinear problem. To change it to a linear problem one must introduce an intermediate matrix or vector, say **M**, which must have more free variables than **R** and **T**. Thus the solution from **M** to **R** and **T** is overdetermined. This is not a problem if everything is accurate since we can always find a consistent decomposition of **M** into **R** and **T**. However in case of noise the "overdeterminedness" from **M** to **R** and **T** is not an asset but a liability. This is because of that for a matrix **M** of more free variables to be consistently decomposed into **R** and **T** of less free variables **M** must satisfy a necessary and surfficient condition. But **M** is sovled from the least square solution of the motion equations for many correspondences of all kinds of errors. Hence the unconstrained least square solution of **M** cannot guarantee a consistent decomposition into **R** and **T**. This problem is discussed in section 6 through an example in detail.

Another fact well known to most motion researchers is that at least 5 point correspondences are needed to get a finite number of motion solutions. However, by introducing some other physical constraints which are available from the correspondences, we argue 4 point correspondences in two views might decide the motion to a finite number of solutions. We shall give more independent equations than unknowns in four point correspondence problem. This will be discussed in detail in section 3. Since by reducing a point correspondence we greatly reduce the dimension of the nonlinear equations, thus a nonlinear solution from 4 point correspondences seems much more feasible than from the traditional 5 point equations. The physical constraints we introduce can not only help to solve the motion but also to judge the motion estimation. That is, all the physical constraints can be used to solve the motion as well as to serve the performance criterion for a globally optimal solution in case multiple point correspondences are available.

We also present some basic theorems for plane and general surface motion. Based on these theorems and those established earlier by others we suggest some algorithms for plane and general surface motion solution. However the emphasis of this paper is put on general motion solution from multiple point correspondences because this problem is essential. Almost every theorem presented in this paper has been verified by experiment. Our discussion here is mainly on monocular vision because it is also essential in motion analysis. But we presents some theoretical results for binocular vision.

Our solution of general surface motion is based on a planar model and thus only requires three correspondences by searching the fourth in a reasonable window for motion parameter estimation. But at least one more correspondence is needed for the judgement f the optimal solution. Several linear algorithms with 8-point correspondences were proposed by Zhuang ([22],1986), Longuet-Higgins ([23][24],1981,1984) and Weng ([27], 1987), for general surface motion solution. However, since these algorithms use as many as 8 point coprrespondences, they seem less robust and have some problems in practice. First, the 8 points have to satisfy a surface assumption(Zhuang,[22]). But in practice, how can one assure that he can always find such correspondences? Second, to make the solution as robust as possible, these 8 points should separate as sparsely as possible. This requirement may cause problem in multiple object cases becasue one doesn't know whether all these points belong to the same object before he solves the motion. Our method may relatively ease the problem because our solution needs fewer points. Third, these algorithms are two step algorithms and hence have the problems discussed above. At last, they cannot deal with the pure rotation case and even small translation case in a unified form. This is because they do not have a unified criterion to judge which solution is best: pure rotation, or a rotation plus a translation. Problems with traditional 8-point linear algorithms are discussed in section 6.

It sounds strange that a planar model can solve general surface motion since a planar model requires the 4 point correspondences belong to the same plane. We shall show, however, this requirement will not restrict the usefulness of our method by adopting a 3 point correspondence algorithm in practice. In binocular vision, we shall show, if absolute or relative depths are known, three noncolinear point correspondences suffice to decide the motion uniquely. Since in binocular vision it is not difficult to find three points' correspondences with their depths known, thus this assumption will not block our methed to be applied to binocular vision. In monocular vision case, since we do not assume we have any knowledge of the surfaces, thus we cannot assure that the four points we choose, which are needed for the motion solution, lie in the same plane. However, by playing a small trick and searching in a reasonable window we only need three correspondences to solve the plane motion. Since three noncolinear points in space always define a plane, thus we do not need any knowledge of the surface to find the motion. And from a sequential point of view of the motion problem, we shall show, if the motion with a nonzero translation component between two image frames is known, we can always find the plane defined by any 3 noncolinear points in both frames if only we know their correspondences. And if we know the plane that the 3 points lie and their correspondences, then we can get up to four sets of motion solution. With the help of other correspondences we can make the solution unique.

The plane model solution may also have direct usages in many applications such as navigation, airborn camera motion analysis, industry robotics or whenever plane model suits or plane information is available.

So far few criterions are proposed to serve the judge of the correctness of the motion solution. The existing criteria like the least square are just used to solve the motion. By these criteria one can not guarantee that each correspondence plays a equal role in the performance index. Some other criteria incorporate in the surface smoothness constraint and thus can not distinguish noise from surface discontinuity. In this paper we shall propose another criterion for optimal motion solution. This new proposed criterion is an equalized one in the sense that every correspondence plays an equal role in the judgement of the optimal solution. But the computation of it takes more time than the least square criterion.

The accuracy of any motion solution algorithm depends on that of correspondence. If the correspondences are wrong, then the motion solution will be wrong either. While the motion is known, we would like to point out that the correspondence of a point will lie in a *motion epipolar line*. After we solve one motion somehow, we can use the motion epipolar line constraint to find motion bondaries and hence solve other motions. With the motion parameters solved, we can also improve the accuracy of correspondences we get earlier by other methods since the matching problem is now a one dimension problem. If the translation is not zero, then we can get an estimation of the surface shape. Iteratively we can improve our understanding of motion and the surface at the same time.

To avoid the camera calibration problem and to judge the robustness of the algorithms themselves, we only carry on simulation experiments. We take a real picture with a real camera and assign an artificial surface to the scene. We assume the image coordinates be exact. Then we simulate a motion via the perfect transformation model. Then the motion is exactly known, but the correspondences are known to within 0.5 pixel which is the quantization error. Then with this correspondence precision we recover the motion given a number of correspondences. We compare the results of our algorithm with the 8-point linear algorithm. With the results from matching we also give estimations of motions.

In section 2 we give two representations and some basic theorems of motion and the basic notations in this paper. In section 3 we propose two physical constraints other than the rigidity constraint and show that 4 point correspondences in two views might decide the motion to a finite number of solutions. In section 4 we give a closed form solution for plane motion from an approximate approach. In section 5 we give some basic theorems and an algorithm for the plane motion solution from the perfect transformation approach. This algorithm is also used in our general surface motion solution. In section 6 we discuss the problems of the 8-point algorithm and give our algorithm under this methodology. In section 7 we propose our criterion for optimal motion estaimation. In section 8 we present our algorithms for general motion solution with multiple point correspondences. In section 9 we discuss the relationship between the long range and short range motion. In section 10 we give some experiment results. Finally in section 11 we summarize our conclusions and suggestions.

Our results presented here are just part of our ongoing work. We shall issue our results in a series of papers. In this paper we only discuss perspective projection.

2. Representations of and Basic Theorems for Motion

In the literature there are two most commonly used representations of motion: velocity decomposition model and perfect transformation model. We will discuss these two representations one by one. The first model is only valid for small rotation, but the second is valid for any rigid motion. It seems to us the second model gives much more robust solution even in the small motion cases, though the solution is usually more difficult.

We shall use the coordinate system shown in Fig. 2-1, where $x-y$ denotes image coordinates and $X-Y-Z$ denotes real world coordinates, and $\overline{of}$ = f is the focal length of the camera. Without losing generality we assume f = 1 in this paper. In case $f \neq 1$, one should replace x by x/f and y by y/f everywhere. Thus an image point (x,y) represents the projection of an scene point (X,Y,Z), and this relation is denoted by

$$\mathbf{P}: \; (X, Y, Z) \rightarrow (x, y)$$

In our motion representations we assume that the camera be static, and hence all the motion parameters represent the object motion relative to the camera. Through this paper, we shall use following notations frequently. We first give the rule here. Bold capitals represent vectors or matrices, capitals represent coordinates in the space, lowercase letters represents coordinates in the image plane or elements of vectors or matrices. Unless specified, (x,y) is always the projection of (X,Y,Z), and (x_i,y_i), (x'_i,y'_i) are the projections of (X_i,Y_i,Z_i), (X'_i,Y'_i,Z'_i) separately, where i is any subscript. And whenever appeared, $\mathbf{\Theta}$, $\mathbf{\Theta}_i$, $\mathbf{\Theta}'_i$ always denotes $[\,x\ y\ 1\,]^T$, $[\,x_i\ y_i\ 1\,]^T$, $[\,x'_i\ y'_i\ 1\,]^T$. And unless specified we always assume a coordinate with prime correspond to the coordinate without prime. Thus

$$X' \approx X + \dot{X}, \; x' \approx x + \dot{x}$$

$$\mathbf{X} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = Z\mathbf{\Theta} = Z\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \quad \mathbf{X}'_i = \begin{bmatrix} X'_i \\ Y'_i \\ Z'_i \end{bmatrix} = Z'_i\mathbf{\Theta}'_i = Z'_i\begin{bmatrix} x'_i \\ y'_i \\ 1 \end{bmatrix}$$

and so on. And we use $\longleftrightarrow$ to denote "corresponding" between two image frames. That is, $(X, Y, Z) \longleftrightarrow (X', Y', Z')$ means (X, Y, Z) coresponds to (X', Y', Z') in the space, and $(x, y) \longleftrightarrow (x', y')$ means (x, y) corresponds to (x', y') in the image plane. Sometimes it would be helpful to treat a point in space as a vector and vice versa. So in this paper we may alternatively treat a point in space as a vector or treat a vector as a point in the space.

The above denotion is meaningful because from similar triangles in Fig. 2-1, we get

$$x = f\,\frac{X}{Z} = \frac{X}{Z} \qquad [2\text{-}1]$$

$$y = f\,\frac{Y}{Z} = \frac{Y}{Z} \qquad [2\text{-}2]$$

and

$$\dot{x} = \frac{\dot{X}}{Z} - \frac{X\dot{Z}}{Z^2} = \frac{\dot{X}}{Z} - \frac{x\dot{Z}}{Z} \qquad [2\text{-}3]$$

$$\dot{y} = \frac{\dot{Y}}{Z} - \frac{Y\dot{Z}}{Z^2} = \frac{\dot{Y}}{Z} - \frac{y\dot{Z}}{Z} \qquad [2\text{-}4]$$

2.1 Velocity Decompostion Model

As is well known, any 3-D rigid movement can be decomposed into two components ([5]):

1. rotation with velocity $\vec{\omega}$ around a center $\mathbf{C}$ $(X_{c,}\ Y_{c,}\ Z_c)$, the projection of which is c $(x_{c,}\ y_c)$
2. translation with velocity $\vec{V}_c$.

The selection of $\mathbf{C}$, which can be any point, will be discussed later. Let $\hat{i}, \hat{j}, \hat{k}$ be the unit vectors along X, Y, Z axis respectively, and let

$$\vec{r}_c = X_c\,\hat{i} + Y_c\,\hat{j} + Z_c\,\hat{k} \qquad [2\text{-}5]$$

$$\vec{r} = X\,\hat{i} + Y\,\hat{j} + Z\,\hat{k} \qquad [2\text{-}6]$$

$$\rho = \vec{r} - \vec{r}_c = \rho_X\,\hat{i} + \rho_Y\,\hat{j} + \rho_Z\,\hat{k} \qquad [2\text{-}7]$$

where

$$\begin{aligned} \rho_X &= X - X_c \\ \rho_Y &= Y - Y_c \\ \rho_Z &= Z - Z_c \end{aligned} \qquad [2\text{-}8]$$

and

$$\rho_X^2 + \rho_Y^2 + \rho_Z^2 = \rho^2 = \text{constant} \qquad [2\text{-}9]$$

by the rigidity condition.

Taking derivatives of $\vec{r}$ and $\vec{r}_c$ over time,

$$\vec{V}_c = \dot{\vec{r}}_c = \dot{X}_c\,\hat{i} + \dot{Y}_c\,\hat{j} + \dot{Z}_c\,\hat{k} \qquad [2\text{-}10]$$

$$\dot{\vec{r}} = \dot{X}\,\hat{i} + \dot{Y}\,\hat{j} + \dot{Z}\,\hat{k} \qquad [2\text{-}11]$$

and representing $\vec{\omega}$ as

$$\vec{\omega} = \omega_X\,\hat{i} + \omega_Y\,\hat{j} + \omega_Z\,\hat{k} \qquad [2\text{-}12]$$

then, the motion of an object point is given by

$$\begin{bmatrix} \dot{X} \\ \dot{Y} \\ \dot{Z} \end{bmatrix} = \begin{bmatrix} \dot{X}_c \\ \dot{Y}_c \\ \dot{Z}_c \end{bmatrix} + \begin{bmatrix} \omega_X \\ \omega_Y \\ \omega_Z \end{bmatrix} \times \begin{bmatrix} \rho_X \\ \rho_Y \\ \rho_Z \end{bmatrix} = \begin{bmatrix} \dot{X}_c + \rho_Z\,\omega_Y - \rho_y\,\omega_Z \\ \dot{Y}_c + \rho_X\,\omega_Z - \rho_Z\,\omega_X \\ \dot{Z}_c + \rho_Y\,\omega_X - \rho_X\,\omega_Y \end{bmatrix} \qquad [2\text{-}13]$$

Differentiating [2-9], we get

$$(X - X_c)\,(\dot{X} - \dot{X}_c) + (Y - Y_c)\,(\dot{Y} - \dot{Y}_c) + (Z - Z_c)\,(\dot{Z} - \dot{Z}_c) = 0. \qquad [2\text{-}14]$$

Then, from [2-5], [2-6], and [2-13], we have

$$\dot{x} = \frac{1}{Z}\,(\dot{X}_c + \rho_Z\,\omega_Y - \rho_Y\,\omega_Z) + \frac{x}{Z}\,(\dot{Z}_c - \rho_Y\,\omega_X + \rho_X\,\omega_Y)$$

$$\dot{y} = \frac{1}{Z}\,(\dot{Y}_c + \rho_X\,\omega_Z - \rho_Z\,\omega_X) + \frac{y}{Z}\,(\dot{Z}_c - \rho_Y\,\omega_X + \rho_X\,\omega_Y) \qquad [2\text{-}15]$$

To simplify the representation, we choose

$$X_c = Y_c = Z_c = 0 \qquad [2\text{-}16]$$

Then [2-15] becomes

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} \frac{1}{Z} & 0 & -\frac{x}{Z} & -xy & 1+x^2 & -y \\ 0 & \frac{1}{Z} & -\frac{y}{Z} & -1-y^2 & xy & x \end{bmatrix} \begin{bmatrix} \dot{X}_c \\ \dot{Y}_c \\ \dot{Z}_c \\ \omega_X \\ \omega_Y \\ \omega_Z \end{bmatrix} \qquad [2\text{-}17]$$

2.2 Perfect Transformation Model

Another useful model to represent motion is the perfect transformation one. Suppose $\mathbf{X} = (X\ Y\ Z)^T$ and $\mathbf{X}' = (X'\ Y'\ Z')^T$ are the two positions of the same scene point P at time t and t+Δt. Without losing generality we assume $\Delta t = 1$ in this paper unless specifyed. And assume

$$\mathbf{P}:\ (X, Y, Z) \rightarrow (x, y)$$

$$\mathbf{P}:\ (X', Y', Z') \rightarrow (x', y')$$

Then it is also well-known that $\mathbf{X}$ and $\mathbf{X}'$ are related by a perfect mathematical model[18]

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \mathbf{R} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \mathbf{T} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} \quad [2\text{-}18]$$

where $\mathbf{R}$ is the rotation matrix and $\mathbf{T}$ is the translation vector. $\mathbf{R}$, as is well known, is an orthonormal matrix of first kind and is independent of the selection of the origin of the coordinate system. There are two forms of representations of $\mathbf{R}$ in the literature. One is

$$\mathbf{R} = \begin{bmatrix} n_1^2 + (1-n_1^2)\cos\theta & n_1n_2(1-\cos\theta) - n_3\sin\theta & n_1n_3(1-\cos\theta) + n_2\sin\theta \\ n_1n_2(1-\cos\theta) + n_3\sin\theta & n_2^2 + (1-n_2^2)\cos\theta & n_2n_3(1-\cos\theta) - n_1\sin\theta \\ n_1n_3(1-\cos\theta) - n_2\sin\theta & n_2n_3(1-\cos\theta) + n_1\sin\theta & n_3^2 + (1-n_3^2)\cos\theta \end{bmatrix} \quad [2\text{-}19]$$

where (n_1,n_2,n_3) is the unit direction vector of the rotation axis of θ through the origin of the coordinate system, with

$$n_1^2 + n_2^2 + n_3^2 = 1 \quad [2\text{-}20]$$

The other is

$$\mathbf{R} = \mathbf{A}_X \mathbf{A}_Y \mathbf{A}_Z \quad [2\text{-}21]$$

where

$$\mathbf{A}_X = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_X & -\sin\theta_X \\ 0 & \sin\theta_X & \cos\theta_X \end{bmatrix} \quad [2\text{-}22]$$

$$\mathbf{A}_Y = \begin{bmatrix} \cos\theta_Y & 0 & \sin\theta_Y \\ 0 & 1 & 0 \\ -\sin\theta_Y & 0 & \cos\theta_Y \end{bmatrix} \quad [2\text{-}23]$$

$$\mathbf{A}_Z = \begin{bmatrix} \cos\theta_Z & -\sin\theta_Z & 0 \\ \sin\theta_Z & \cos\theta_Z & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad [2\text{-}24]$$

and the positions of $\mathbf{A}_X$, $\mathbf{A}_Y$ and $\mathbf{A}_Z$ in [2-21] are interchangeable.

One thing needs to be mentioned here is that a rotation of the object around a line not passing the origin of the coordinate system will not be a pure rotation in the representation [2-18]. That is the $\mathbf{T}$ will not be zero in [2-18] in this case. Only a rotation around a line passing the origin will be a pure rotation in this representation.

Besides the orthonormality $\mathbf{R}$ still has another very useful property unnoticed by many researchers. Here we summarize all its properties into a Fact. In our papers we shall give several Facts without proof because they are obvious and familiar to many people. Fact 2.1 can be proved by the rigidity condition.

Fact 2.1 .

Any rotation matrix $\mathbf{R}$ has following properties:

1. Independence from coordinate system: $\mathbf{R}$ is independent of the selection of the coordinate origin.

2. Orthonormality:

$$\mathbf{R}^T\mathbf{R} = \mathbf{I}, \text{ or } \mathbf{R}^T = \mathbf{R}^{-1}, \quad [2\text{-}25]$$

where $\mathbf{I}$ is the identity matrix. Thus

$$r_{ij} = \mathbf{M}_{ij} \quad [2\text{-}26]$$

where $\mathbf{M}_{ij}$ is the algebraic minor determinant of r_{ij}. For example,

$$r_{33} = \mathbf{M}_{33} = r_{11}\, r_{22} - r_{12}\, r_{21}$$

3. Unit determinant:

$$\det(\mathbf{R}) = 1 \qquad [2\text{-}27]$$

4. Rigidity: Given any two vectors $\mathbf{X}_1$and $\mathbf{X}_2$,

$$\mathbf{R}(\mathbf{X}_1 \times \mathbf{X}_2) = (\mathbf{R}\mathbf{X}_1) \times (\mathbf{R}\mathbf{X}_2) \qquad [2\text{-}28]$$

We should note here that our definition of rigidity is different from but equivalent to the traditional definition of rigidity. In fact the traditional rigidity is equivalent to the orthonormality plus the unit determinant property we define here. So unless specified, when we talk about rigidity we always refer to our definition. The reason we use this definition is, as we shall prove in next section, equation [2-28] is a more concise definition of the rigid motion. Property 4 will imply [2-25] and [2-27], and [2-25] and [2-27] are just the requirement of a rigid motion, thus our definition will be equivalent to the traditional rigidity definition. This is to be shown in the next section.

The last property, i.e. rigidity, is extreamely useful but seldom used so far. The procedure of its proof is the same as we use in the proof of Theorem 5.6. The geometrical meaning is obvious if we consider the rotation of a triad. We know that a triad is still a triad and will keep the same chirality structure after rotation. The readers will find [2-28] is extensively used in our work because it not only makes the calculation or representation simpler but also introduces more equations.

From Fact 2.1 we immediately have following theorems:

Theorem 2.1 .

Two unparallel unit vector correspondences in a pure rotation case suffice to decide the rotation matrix uniquely.

Proof: let $\mathbf{V}_i$, i = 1,2, be two unparallel unit vectors and $\mathbf{V}'_i$, i = 1,2, be their correspondences. Thus we have

$$\mathbf{V}'_i = \mathbf{R}\mathbf{V}_i,\ i=1,2 \qquad [2\text{-}29]$$

From Fact 2.1, we should also have

$$\mathbf{V}'_1 \times \mathbf{V}'_2 = (\mathbf{R}\mathbf{V}_1) \times (\mathbf{R}\mathbf{V}_2) = \mathbf{R}(\mathbf{V}_1 \times \mathbf{V}_2)$$

Hence we know $\mathbf{R}$ can be recovered from

$$\mathbf{R} = [\,\mathbf{V}'_1\ \mathbf{V}'_2\ (\mathbf{V}'_1 \times \mathbf{V}'_2)\,][\,\mathbf{V}_1\ \mathbf{V}_2\ (\mathbf{V}_1 \times \mathbf{V}_2)\,]^{-1} \qquad [2\text{-}30]$$

The inversion in [2-30] exists because $\mathbf{V}_1$ and $\mathbf{V}_2$ are not parallel to each other. Q.E.D..

This theorem is frequently used in our work and makes the calcualtion of $\mathbf{R}$ very simple. Unlike many other solutions for $\mathbf{R}$, this solution guarantees the orthonormality of $\mathbf{R}$. We know that one vector correspondence cannot decide the rotation matrix uniquely, hence this solution is the simplest linear algorithm to solve $\mathbf{R}$ from the fewest vector correspondences so far. Immediately we have the following theorems.

Theorem 2.2 .

If $\mathbf{T} = 0$, then two point correspondences without the depths known suffice to decide $\mathbf{R}$ uniquely.

Proof: Let $\mathbf{X}_i$ and $\mathbf{X}'_i$, i = 1,2, be the two correspondence pairs, then $\mathbf{X}'_i = \mathbf{R}\mathbf{X}_i$, i = 1,2, or

$$Z'_i\,\mathbf{\Theta}'_i = \mathbf{R}\;\mathbf{\Theta}_i\,Z_i, \quad i = 1,2 \qquad [2\text{-}31]$$

Take the norm of both sides of [2-31] and use [2-25] and the fact that Z_i, Z'_i are both positive, we get

$$\gamma_i \stackrel{\Delta}{=} \frac{Z'_i}{Z_i} = \frac{\|\mathbf{\Theta}_i\|}{\|\mathbf{\Theta}'_i\|} = \sqrt{\frac{x_i^2 + y_i^2 + 1}{x_i'^2 + y_i'^2 + 1}}\,, \quad i = 1, 2 \qquad [2\text{-}32]$$

Thus, from Theorem 2.1 we know

$$\mathbf{R} \;=\; \Big[\;\gamma_1\mathbf{\Theta}'_1\;\gamma_2\mathbf{\Theta}'_2\;(\gamma_1\mathbf{\Theta}'_1)\times(\gamma_2\mathbf{\Theta}'_2)\Big]\;\Big[\;\mathbf{\Theta}_1\;\mathbf{\Theta}_2\;\mathbf{\Theta}_1\times\mathbf{\Theta}_2\Big]^{-1} \qquad [2\text{-}33]$$

Since by two correspondences we mean $(x_1, y_1) \neq (x_2, y_2)$, hence $\mathbf{\Theta}_1$ is always not parallel to $\mathbf{\Theta}_2$. This guarantees the existence of the inversion on the right side of [2-33]. Q.E.D..

Note that in the case of a pure rotation, if the two correspondences are correct, then, besides [2-31], from the orthonormality of $\mathbf{R}$ we must also have

$$\gamma_1\gamma_2\mathbf{\Theta}'_1\cdot\mathbf{\Theta}'_2 \;=\; \mathbf{\Theta}_1\cdot\mathbf{\Theta}_2 \qquad [2\text{-}34]$$

[2-34] is a necessary condition that a pure rotation occurs. One should note that the premise in Theorem 2.2 is that we have already known the translation is zero. If the translation is not zero, then even an infinite number of correspondences may not decide the motion uniquely if they all align.

Theorem 2.3 .

If depths are known, three noncolinear space point correspondences suffice to decide the motion uniquely.

Proof: Let $\mathbf{X}_i$ and $\mathbf{X}'_i$, i = 1,2,3, be the three correspondence pairs, then $\mathbf{X}'_i = \mathbf{R}\mathbf{X}_i + \mathbf{T}$, i = 1,2,3, or

$$\mathbf{X}'_i - \mathbf{X}'_1 = \mathbf{R}\ (\mathbf{X}_i - \mathbf{X}_1), \quad i = 2,3 \qquad [2\text{-}35]$$

Since $\mathbf{X}_i$, i = 1,2,3, are not colinear, thus $\mathbf{X}_2 - \mathbf{X}_1$ is not parallel to $\mathbf{X}_3 - \mathbf{X}_1$, then from Theorem 2.1 we know $\mathbf{R}$ can be uniquely decided by

$$\mathbf{R} = \left[\mathbf{X}'_2 - \mathbf{X}'_1\ \mathbf{X}'_3 - \mathbf{X}'_1\ (\mathbf{X}'_2 - \mathbf{X}'_1) \times (\mathbf{X}'_3 - \mathbf{X}'_1)\right] \left[\mathbf{X}_2 - \mathbf{X}_1\ \mathbf{X}_3 - \mathbf{X}_1\ (\mathbf{X}_2 - \mathbf{X}_1) \times (\mathbf{X}_3 - \mathbf{X}_1)\right]^{-1} \qquad [2\text{-}36]$$

and $\mathbf{T}$ can be decided by

$$\mathbf{T} = \mathbf{X}'_i - \mathbf{R}\mathbf{X}_i, \quad i = 1,2,3 \qquad [2\text{-}37]$$

Q.E.D..

If depths are only known in one image frame, then we may have up to four sets of motion solution for the perfect transformation model. But from [2-17] we know there is a unique solution for the approximation model. So if the motion is small we can use both representations to get a unique solution. This case is discussed in section 5 in detail. Still we have another fact:

Fact 2.2 .

1. If $\mathbf{T} = 0$, and $\mathbf{R}$ is known, then the correspondence (x', y') of (x,y) is uniquely decided by

$$x' = \frac{r_{11}\, x + r_{12}\, y + r_{13}}{r_{31}\, x + r_{32}\, y + r_{33}} \qquad [2\text{-}38]$$

$$y' = \frac{r_{21}\, x + r_{22}\, y + r_{23}}{r_{31}\, x + r_{32}\, y + r_{33}}$$

2. If $\mathbf{R}$ is known and and $\mathbf{T} \neq 0$ is known up to a scalar, then the correspondence (x', y') of (x,y) lies in a *motion epipolar line* defined by

$$[\ x'\ y'\ 1\ ]\ (\mathbf{T} \times \mathbf{R})\ [\ x\ y\ 1\ ]^T = 0 \qquad [2\text{-}39]$$

After $\mathbf{R}$ is solved somehow, then with how many point correspondences and in what condition can one uniquely decide the translation up to a scalar? The following theorem answers this question. Before we state the theorem we first introduce a two step representation equivalent to [2-18] for the perfect transformation model as following:

$$\mathbf{X}'' = \mathbf{R}\,\mathbf{X}$$

$$\mathbf{X}' = \mathbf{X}'' + \mathbf{T} \qquad [2\text{-}40]$$

And we see if

$$\mathbf{P}:\ (X'',Y'',Z'') \to (x'',y'')$$

then equation [2-39] is equivalent to

$$[\,x'\ y'\ 1\,]\,(\mathbf{T} \times [\,x''\ y''\ 1\,]^{T}) = 0,\ \text{or} \qquad [2\text{-}41]$$

$$[\,x'\ y'\ 1\,]\begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{bmatrix}\begin{bmatrix} x'' \\ y'' \\ 1 \end{bmatrix} = 0,\ \text{or} \qquad [2\text{-}42]$$

$$(y''-y')\,t_1 + (x'-x'')\,t_2 + (y'x''-x'y'')\,t_3 = 0; \qquad [2\text{-}43]$$

From Fact 2.2 we know if $\mathbf{R}$ is decided, then (x'', y'') is uniquely decided by (x, y). So what we need to show is the condition under which [2-43] has exactly one independent solution. If $\mathbf{T} = 0$ we see for all points in the image we must have $x' = x''$ and $y' = y''$. Now let's consider the case $\mathbf{T} \neq 0$. From the projection law [2-1] and [2-2] we have

$$x' = \frac{X'' + t_1}{Z'' + t_3},\ y' = \frac{Y'' + t_2}{Z'' + t_3} \qquad [2\text{-}44]$$

$$x'' = \frac{X''}{Z''},\ y'' = \frac{Y''}{Z''} \qquad [2\text{-}45]$$

To make $x' = x''$ we must have

$$x''t_3 = t_1 \qquad [2\text{-}46]$$

and to make $y' = y''$ we must have

$$y''t_3 = t_2 \qquad [2\text{-}47]$$

Because of [2-46] and [2-47], for a given translation, except those points of infinite depths, there is at most one point such that $x' = x''$ and $y' = y''$ hold at the same time. Now we can state the theorem.

Theorem 2.4 .

If $\mathbf{R}$ is known somehow then $\mathbf{T}$ can be decided to within a scalar by two image points (x_i, y_i), $i = 1,2$, of finite depths and their correspondences (x'_i, y'_i), $i = 1,2$, by solving [2-43] iff

1.

$$(x'_i, y'_i) = (x''_i, y''_i), \ \ i= 1,2 \qquad [2\text{-}48]$$

where (x''_i, y''_i), $i = 1,2$, are calculated according to [2-38]. In this case $\mathbf{T} = 0$, or

2.

$$(x'_i, y'_i) = (x''_i, y''_i), \text{ but } (x'_j, y'_j) \neq (x''_j, y''_j), \ i{\neq}j, \ \ i,j{=}1,2 \qquad [2\text{-}49]$$

in this case $\mathbf{T} \neq 0$, or

3.

$$(x'_i, y'_i) \neq (x''_i, y''_i), \ \ i= 1,2 \qquad [2\text{-}50]$$

and (x'_1, y'_1), (x'_2, y'_2), (x''_1, y''_1) and (x''_2, y''_2) are not colinear in the image plane. In this case $\mathbf{T} \neq 0$.

Proof: Part 1 is true because for points of finite depths if and only if $\mathbf{T} = 0$ then we have $x'_i = x''_i$ and $y'_i = y''_i$, $i = 1,2$, at the same time, or equivalently we have condition [2-48].

Now let's prove Part 2. If [2-48] holds for one correspondence pair, say for i = 1, but does not hold for the other correpondence pair, say for i = 2, then we can conclude that $\mathbf{T} \neq 0$. Otherwise [2-48] should hold for both pairs. Then from [2-46] and [2-47] we know $\mathbf{T}$ can be solved up to a scalar from

$$t_1 = x''_1 t_3, \ \ t_2 = y''_1 t_3 \qquad [2\text{-}51]$$

where $t_3 \neq 0$.

In Part 3 two pairs of correspondences will give two equations as

$$\begin{bmatrix} y''_1-y'_1 & x'_1-x''_1 & y'_1x''_1-x'_1y''_1 \\ y''_2-y'_2 & x'_2-x''_2 & y'_2x''_2-x'_2y''_2 \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} \triangleq \mathbf{A}\,\mathbf{T} = 0 \qquad [2\text{-}52]$$

We know that $\mathbf{T}$ must have at least one independent solution. If $\mathbf{T}$ is to be uniquely decided up to a scalar by two pairs of correspondences then the matrix $\mathbf{A}$ constructed by them must have a rank 2. Since [2-48] does not hold, thus $\mathbf{A}$ has rank 1 iff there exists a non-zero number α such that

$$[y''_2-y'_2 \;\; x'_2-x''_2 \;\; y'_2x''_2-x'_2y''_2] = \alpha[y''_1-y'_1 \;\; x'_1-x''_1 \;\; y'_1x''_1-x'_1y''_1] \qquad [2\text{-}53]$$

or equivalently

$$\frac{y'_2-y''_2}{x'_2-x''_2} = \frac{y'_1-y''_1}{x'_1-x''_1}, \text{ and } \frac{y'_2-y'_1}{x'_2-x'_1} = \frac{y'_1-y''_1}{x'_1-x''_1} \qquad [2\text{-}54]$$

[2-54] is just the colinearity condition for the four points (x'_1,y'_1), (x'_2,y'_2), (x''_1,y''_1),(x''_2,y''_2). Hence iff (x'_1,y'_1), (x'_2,y'_2), (x''_1,y''_1),(x''_2,y''_2) are not colinear then **T** has only one independent solution.

Note that these three situations cannot happen at the same time and at least one situation must happen for **T** to be recovered from two pairs of correspondences. Thus the conditions we list in Part 1, 2 and 3 are necessary and sufficient for the translation vector **T** to be recovered from two correspondences. Q.E.D..

Equation [2-39] means if the motion is known but not the depth, then the search of the correspondence is a one dimensional problem. This gives us other choices for stereo camera configuration. The currently most used configuration of the stereo pair of cameras requires the two image planes be coplanar. This requirement, though ease the representation of the epipolar lines, yet has some problems. First, it cannot guarantee the cameras have the largest common vision field when the span of the cameras is fixed. Second, it requires a high precision of installation. Third, small error in measuring the positions and orientations of the cameras may cause a large error of the depth estimation. And last, it requires both cameras always have the same motion or, at most, the two cameras can only slide along their installation line. We find this requirement is not necessary if we can measure the relative position of the two cameras somehow.

The following fact tells us that depth information is directly available from motion except for one point if the translation is not zero.

Fact 2.3 .

If **R** is known and $\mathbf{T} = \alpha[\, t_1 \; t_2 \; t_3 \,]^T \neq 0$ is known up to a scalar α, then the depth Z of the scene point corresponding to a image point (x,y) can be decided up to a scalar by the positions of its correspondence (x',y') and itself through

$$Z = \alpha \frac{t_1 - x' t_3}{x(r_{31}\, x' - r_{11}) + y(r_{32}\, x' - r_{12}) + r_{33}\, x' - r_{13}}, \quad \text{or}$$

$$Z = \alpha \frac{t_2 - y' t_3}{x(r_{31}\, y' - r_{21}) + y(r_{32}\, y' - r_{22}) + r_{33}\, y' - r_{23}} \qquad [2\text{-}55]$$

where the sign of α is to be chosen to make the depth positive, and the relative depth of the scene point corresponding to (x', y') is given by

$$Z' = (r_{31}\, x + r_{32}\, y + r_{33})Z + t_3 \qquad [2\text{-}56]$$

unless in the rare case

$$t_1 = x' t_3, \quad \text{and} \quad t_2 = y' t_3 \qquad [2\text{-}57]$$

For the point (x_s, y_s) satisfying [2-57] we can use the rigidity equations to solve the depth indirectly:

$$\| Z_s \Theta_s - Z_i \Theta_i \|^2 = \| Z'_s \Theta'_s - Z'_i \Theta'_i \|^2, \quad i = ,1, \cdots ,n \qquad [2\text{-}58]$$

At least two auxiliary points are needed to solve the depth for (x_s, y_s). ∎

In the above Fact, we need only to know one of x' and y' to decide the depth. In the correspondence process, some times y' or sometimes x' is uncertain. But we only need one of them to estimate the depth. We shall discuss this fact in more detail in section 3.

2.3 The Relation between Two Representations

Both the two representations listed here describe any 3-D rigid motion. However, the velocity decomposition model can only be valid for a small rotation while the perfect transformation model is valid for any rigid motion. In the case the the rotation is small, for example, ω_X, ω_Y, ω_Z are all smaller than 10 degrees, these two representations are related through the following fact:

Fact 2.4 .

If the coordinate systems of both representations are chosen the same and the motion center is chosen as the coordinate origin, i.e., [2-16] holds, then

$$t_1 = \dot{X}_c, \; t_2 = \dot{Y}_c, \; t_3 = \dot{Z}_c, \; r_{11} \approx r_{22} \approx r_{33} \approx 1$$

$$\omega_X \approx \theta_X \approx r_{32} \approx -r_{23}, \quad \omega_Y \approx \theta_Y \approx r_{13} \approx -r_{31}, \quad \omega_Z \approx \theta_Z \approx r_{21} \approx -r_{12} \qquad [2\text{-}59]$$

3. Motion Constraints and 4-point Nonlinear Equations

So far the motion epipolar line constraint is most often used to solve the motion and serve the performance criterion of optimal estimation of motion. But there are still many other constraints which can serve the same purposes. The reason for this is that its representation is linear and only involves motion parameters and image plane coordinates and hence does not contain the unknown depths. However in section 6 we shall show that the two step linear solution based on the motion epipolar line equation will generally not work. And our algorithm to be introduced in section 8 needs to search in a reasonable space and hence is also intrinsically nonlinear. Now that we cannot get a robust closed form solution for the motion problem why don't we directly solve the nonlinear equations? So far few results have been reported in this direction. The problem may be in the difficulty of the solution of high dimensional nonlinear polynomial equations. And we still do not have an effective method to integrate multiple measurements to get an optimal solution for nonlinear equations in the sense of least square or something equivalent. In this section we only explore the feasibility of the nonlinear motion equations. We shall prove four noncolinear point correspondences might determine the motion to a finite number of solutions and hence we can greatly reduce the dimension of the nonlinear equations.

It is long believed that at least five correspondences are needed to give a finite number of motion solutions if the corresponding scene points lie on a general surface. For example if one wants to solve the motion parameters from the motion epipolar line equation [2-39] in a nonlinear way, then he is believed to need at least 5 point correspondences to get five independent equations. We see [2-39] only involves the motion parameters as unknowns. Another sets of nonlinear equations only involving depths as unknowns are got from the the *traditional rigidity* or *orthonormality* consideration, that is, equation [2-25]. Let's list the equations here. Let $\mathbf{X}'_i = \mathbf{\Theta}'_i Z'_i$ be the correspondence of $\mathbf{X}_i = \mathbf{\Theta}_i Z_i$, $i = 1,2,\cdots,n$. Then from the motion equation [2-18] we have

$$Z'_i \mathbf{\Theta}'_i = Z_i \mathbf{R} \mathbf{\Theta}_i + \mathbf{T}, \; i=1,2,\cdots,n \qquad [3\text{-}1]$$

The motion epipolar line equation is repeated here for convenience:

$$\mathbf{\Theta}'^{T}_i (\mathbf{T} \times \mathbf{R}) \mathbf{\Theta}_i = 0, \; i=1,2,\cdots,n \qquad [3\text{-}2]$$

Three correspondences will give three independent equations from the *orthonormality* constraint. For example, given three pairs of correspondences with indices being 1,2,3, we shall have,

$$\|\Theta'_i Z'_i - \Theta'_1 Z'_1\|^2 = \|\Theta_i Z_i - \Theta_1 Z_1\|^2, \ i = 2,3 \quad [3\text{-}3]$$

$$(\Theta'_2 Z'_2 - \Theta'_1 Z'_1)\cdot(\Theta'_3 Z'_3 - \Theta'_1 Z'_1) = (\Theta_2 Z_2 - \Theta_1 Z_1)\cdot(\Theta_3 Z_3 - \Theta_1 Z_1) \quad [3\text{-}4]$$

Introducing one more correspondence will add in two new unknowns and was believed to be able to bring in only three independent equations as following, for example,

$$\|\Theta'_i Z'_i - \Theta'_4 Z'_4\|^2 = \|\Theta_i Z_i - \Theta_4 Z_4\|^2, \ i = 1,2,3 \quad [3\text{-}5]$$

since the size of the quadrihedron constructed by the four space points has already been decided by them. This is shown in Fig. 3-1. But so far we only have 6 equations containing 8 - 1 = 7 unknown depths. Thus at least 5 correspondences were thought to be needed to give a finite number of solutions for the motion since 5 correspondences were thought to give 9 independent equations and involve 10–1 = 9 unknown depths (we can scale one depth to unit). After the depths are solved then we can solve the motion by Theorem 2.3.

One question we want to ask here now is whether the equations from the motion e pipolar line constraint are dependent on those from orthonormality constraint? If so, how many independent equations in total do we have? If not, it seems we may have more independent equations from four point correspondences, then it seems the motion may be determined by four correspondences. And another question we want to inquire here is: are the motion epipolar constraint and orthornormality constraint what we only have? If not, other constraints may introduce more equations.

In the following when we talk about a point it might be a point in the image or in the scene according to the context, but when we talk about a depth it always means a depth of a scene point corresponding to a image point. Before we begin, we must assume the following condition:

Surface Condition 1.

The points used for correspondences do not all lie on a line in space. ||

otherwise, a rotation around the line they construct can never be decided. To enforce this condition we only need to require the points in the image plane be not all colinear since the projection of a line in space must be a line or a dot in the image plane.

Before we proceed we give some lemmata for convenience of discussion.

Lemma 3.1.

The correspondence $\mathbf{\Theta}'$ of $\mathbf{\Theta}$ is parallel to $\mathbf{R\Theta}$ iff $\mathbf{T} = 0$ or $\mathbf{T} \neq 0$ and $\mathbf{T}$ is parallel to $\mathbf{\Theta}'$ or $\mathbf{R\Theta}$, where $\mathbf{R}$ is the rotation matrix and $\mathbf{T}$ is the translation vector.

Proof: This can be directly seen from [3-1]. Q.E.D..

In Section 2 we have shown if $\mathbf{T} = 0$, then two correspondences uniquely decide the rotation. Thus in this section we assume the translation be nonzero, i.e., $\mathbf{T} \neq 0$. Since the motion epipolar line equations are in terms of the motion parameters and the traditional rigidity equations are in terms of the depths, to decide whether they are dependent or not we need a transformation from depths to motion parameters and the inverse. The following two lammata give us the transformation. From Theorem 2.3 we directly have the following lemma:

Lemma 3.2.

The motion parameters can be represented by the depths of three noncolinear scene points and their correspondences with all the image positions known as following:

$$\mathbf{R} = [\,(\mathbf{\Theta}'_3 Z'_3 - \mathbf{\Theta}'_1 Z'_1)\ (\mathbf{\Theta}'_2 Z'_2 - \mathbf{\Theta}'_1 Z'_1)\ (\mathbf{\Theta}'_3 Z'_3 - \mathbf{\Theta}'_1 Z'_1)\times(\mathbf{\Theta}'_2 Z'_2 - \mathbf{\Theta}'_1 Z'_1)\,]\cdot$$
$$[\,(\mathbf{\Theta}_3 Z_3 - \mathbf{\Theta}_1 Z_1)\ (\mathbf{\Theta}_2 Z_2 - \mathbf{\Theta}_1 Z_1)\ (\mathbf{\Theta}_3 Z_3 - \mathbf{\Theta}_1 Z_1)\times(\mathbf{\Theta}_2 Z_2 - \mathbf{\Theta}_1 Z_1)\,]^{-1} \qquad [3\text{-}6]$$

$$\mathbf{T} = Z'_i \mathbf{\Theta}'_i - Z_i \mathbf{R\Theta}_i, \quad i = 1,2,3 \qquad \| \qquad [3\text{-}7]$$

We can also represent all the depths as functions of the motion $\mathbf{R}$ and $\mathbf{T}$. From [3-1] we have

$$[\,-\mathbf{R\Theta}_i\ \ \mathbf{\Theta}'_i\,]\begin{bmatrix} Z_i \\ Z'_i \end{bmatrix} = \mathbf{T} \qquad [3\text{-}8]$$

The above equation will have an infinite number of solutions for Z'_i and Z_i if $\mathbf{T}$ is parallel to $\mathbf{\Theta}'_i$, and a definite solution

$$\begin{bmatrix} Z_i \\ Z'_i \end{bmatrix} = \begin{bmatrix} \mathbf{\Theta}_i^T\mathbf{\Theta}_i & -\mathbf{\Theta}_i^T\mathbf{R}^T\mathbf{\Theta}'_i \\ -\mathbf{\Theta}_i^T\mathbf{R}^T\mathbf{\Theta}'_i & \mathbf{\Theta}'^T_i\mathbf{\Theta}'_i \end{bmatrix}^{-1} \begin{bmatrix} -\mathbf{\Theta}_i^T\mathbf{R}^T \\ \mathbf{\Theta}'^T_i \end{bmatrix} \mathbf{T}, \quad i=1,\cdots,n \qquad [3\text{-}9]$$

if $\mathbf{T}$ is not parallel to $\mathbf{\Theta}'_i$. Given a rotation $\mathbf{R}$ and a nonzero translation $\mathbf{T}$ there is at most one point in the image plane such that $\mathbf{\Theta}'_i$ is parallel to $\mathbf{T}$. Thus we have the following lemma:

Lemma 3.3.

For a given motion with a nonzero translation **T**, the depths of all surface points but a particular point Θ such that $\mathbf{R}\Theta$ is parallel to **T**, can be directly solved from their projections on one image plane and the correspondences of them in the other image plane from [3-9] without other structural information. ∎

Lemma 3.3 tells us that all but one depths can be represented by the motion parameters and the image plane coordinates. However [3-9] is not the only way to solve [3-8], since [3-8] has three equations but only two unknowns and hence many solution forms are available. Another alternative solution for [3-8] has been given in [2-55] and [2-56]. If the correspondences and the motion parameters are correct all the solutions must be the same. We should point out that [3-9] is not a good form to be used in practice since it complicates the representation. For convenience of discussion, we introduce the following Surface Condition.

Surface Condition 2.

All the depths of the image points used for correspondences can be directly represented by the motion parameters. ∎

Now we have got the transformations we desire. But both the representations in [3-9] and [3-6], [3-7] are too complicate for substitution. So we shall heuristically prove the the orthornomality equations and the motion epipolar equations are not equivalent to each other. One can also replace [3-9] into [3-3] ~ [3-5] or replace [3-6], [3-7] into [3-2] to prove the inequivalence. But we shall prove it by some counter example. Let's assume we only have one correspondence. Then motion epipolar line constraint (MELC) gives us one equation, but traditional rigidity constraint (TRC) gives no equation involving depths only. And if we have two correspondences then MELC gives two equations for five unknowns, and TRC gives only one equation for three unknowns. Now let's consider four correspondences. MELC gives exactly four equations for five unknown motion parameters, but TRC will give 6 independent equations for just 8 - 1 = 7 unknown depths from the traditional rigidity constraint. Now from lemma 3.3 we know if the four point correspondences satisfy Surface Condition 2 then we can express the depths in the 6 rigidity equations in terms of motion parameters. Thus we shall have 6 independent equations for only 5 unknown motion parameters in general! Hence it is obviously the motion should be decided up to a finite number of solutions at least sometimes in the four point correspondence problem. Again let's introduce the fifth correspondence. It was believed at most 3 more independent equations can be brought in by the fifth correspondence. But in fact at most 4 independent equations can be introduced. To see why the fifth correspondence introduce 4 more equations not only 3 equations let's have a look on Fig. 3-2. There we see three equations will still give an ambiguous shape for the polyhedron constructed by the five points and the orthonormality equations. We still need another equation to uniquely decide the shape of the polyhedron uniquely. And from now on everytime one more correspondence is introduced at most 4 more independent orthonormality equations will be introduced! Thus 5 correspondences might give 10 equations for 10 - 1 = 9 unknown depths. But if we represent them in terms of the motion parameters we shall have 10 independent traditional rigidity equations for only five unknown motion parameters. Again we see that five correspondences give exactly five motion epipolar line equations for five unknowns. Hence we have so far shown that the MELC's are not equivalent to TRC's.

Now let's back to the four correspondence problem. Now that MELC is different from TRC why don't we use them at the same time since we only lack one more equation in either constraint equations? We can of course represent the motion parameters in [3-2] by the three noncolinear points and the correspondences to get four more equations for the 7 unknown depths. And also we can represent the depths in [3-3] ~ [3-5] by the motion parameters to get 6 more equations for the 5 unknown motion parametrs! However again we have the difficulty to find out in these equations how many are really indepedent. But obviously we can see that the four correspondences should possibly decide the motion in two view to a finite number of solutions. A physical consideration reveals that the motion epipolar line equations should be implied by the "rigidtiy" equations but not necessary the "traditional rigidty" equations.

We would like to point out that the orthonormality or traditional rigidity equations are not the only equations we can get from the "rigidity" consideration. A rigid motion not only requires the object keep the same shape during motion but also require the orientation and the position of the object be consistent with the motion. Formally we can say that to decide a rigid motion uniquely we need to :

1. Localize a reference point of the object (*reference point constraint*);

2. Keep the shape of the object unchanged (*rigidity constraint*);

3. Make the orientation of the object consistent with the motion (*orientation constraint*).

But so far, the orthonormality equations only have rigidity constraint done. Localizing a reference point will give us three more independent equations and fixing the orientation of the object will give us two more equations. All these 5 equations are not included in the traditional rigid equations and hence must be independent from them in general. However these five equations are generally related to the unknown motion parameters so it seems nothing can be gained from these constraints. Fortunately we have Theorem 2.3 or [3-6] and [3-7]. Now let's find out the 5 more independent equations.

Assume $\mathbf{X}_1$, $\mathbf{X}_2$ and $\mathbf{X}_3$ are three noncolinear points in the space and $\mathbf{R}$ and $\mathbf{T}$ are represented by them and their correspondences in the form of [3-6] and [3-7]. To make sure, assume $\mathbf{T} = Z'_1\Theta'_1 - Z_1\mathbf{R}\Theta_1$. Thus to localize the fourth point, which is selected as the reference point, we have the following equation:

$$Z'_4\Theta'_4 = Z_4\mathbf{R}\Theta_4 + \mathbf{T} = \mathbf{R}(Z_4\Theta_4 - Z_1\Theta_1) + Z'_1\Theta'_1 \qquad [3\text{-}10]$$

However, after the above equation is imposed the below equation contained in [3-5] will be redudent

$$\|\Theta'_1Z'_1 - \Theta'_4Z'_4\|^2 = \|\Theta_1Z_1 - \Theta_4Z_4\|^2 \qquad [3\text{-}11]$$

so we replace it by

$$(\Theta'_1Z'_1 - \Theta'_4Z'_4)\cdot(\Theta'_2Z'_2 - \Theta'_4Z'_4) = (\Theta_1Z_1 - \Theta_4Z_4)\cdot(\Theta_2Z_2 - \Theta_4Z_4) \qquad [3\text{-}12]$$

To fix the orientation of the quadrihedron, we impose the following equation

$$(\Theta'_3Z'_3 - \Theta'_4Z'_4)\times(\Theta'_2Z'_2 - \Theta'_4Z'_4) = \mathbf{R}\,[(\Theta_3Z_3 - \Theta_4Z_4)\times(\Theta_2Z_2 - \Theta_4Z_4)] \qquad [3\text{-}13]$$

[3-13] will give exactly two more independent equations since the angle between $(\Theta'_3Z'_3 - \Theta'_4Z'_4)$ and $(\Theta'_2Z'_2 - \Theta'_4Z'_4)$ has been decided by the traditional rigidity equations. We claim equation [3-4], [3-5], [3-10], [3-12] and [3-13] will give 11 independent equations in general because no one in

them is contained or can be deduced from other equations. Note **R** must be represented by the form of [3-6].

So far we have got 11 independent equations for just 7 unknowns. One may immediately ask: where do these equations come from? Can they be independent? Since it seems all these equations come from the motion representation [3-1], where four correspondences give only 12 equations containing 13 unknowns, and all other equations must be deduced from them, where can these equations come from? So one may conclude they must be dependent. But what we want to ask here is: are the motion representation equations in [3-1] what we only have in a rigid motion? Our answer is no, because a rigid motion should also give us following equations which come from our *rigidity* definition of motion in Fact 2.1:

$$(\Theta'_i Z'_i - \Theta'_j Z'_j) \times (\Theta'_k Z'_k - \Theta'_l Z'_l) = \mathbf{R}\,[(\Theta_i Z_i - \Theta_j Z_j) \times (\Theta_k Z_k - \Theta_l Z_l)], \quad \text{for any } i, j, k, l \qquad [3\text{-}14]$$

All the equations in [3-14] cannnot be covered by the equations in [3-1] because [3-14] is also a basic requirement of a rigid motion. That's why we say our rigidity definition of a rigid motion might be more pertinent to the "rigid" concept. After we impose the equation [3-14] the orthonomality property and the positive determinant property will be redudent and hence we can solve [3-1] and [3-14] togather pretending that we have no knowledge of the rotation matrix **R**. Thus we shall have totally 8 (depth) + 9 (rotation matrix) + 2 (translation vector) = 19 unknowns but as many as 3 * 4 (in [3-1]) + 3 * 6 (in [3-14]) = 30 equations in four correspondences, though some of the equations may not be independent from others.

To show that our rigidity definition implies the orthonormality we first have the following Theorem.

Theorem 3.1 .

Given three noncoplanar vectors $\mathbf{U}_i$, i = 1,2,3, and three vectors $\mathbf{V}_i$, i = 1,2,3, related to the $\mathbf{U}_i$'s through a 3 by 3 matrix **R** via

$$\mathbf{V}_i = \mathbf{R}\mathbf{U}_i, \quad i = 1,2,3 \qquad [3\text{-}15]$$

and

$$\mathbf{V}_i \times \mathbf{V}_j = \mathbf{R}\mathbf{U}_i \times \mathbf{R}\mathbf{U}_j = \mathbf{R}(\mathbf{U}_i \times \mathbf{U}_j), \quad i \neq j, \; i,j = 1,2,3 \qquad [3\text{-}16]$$

Then we must have

$$\mathbf{R}^T\mathbf{R} = \frac{\mathbf{V}_1\cdot(\mathbf{V}_2\times\mathbf{V}_3)}{\mathbf{U}_1\cdot(\mathbf{U}_2\times\mathbf{U}_3)}\mathbf{I} = \alpha\mathbf{I} \qquad [3\text{-}17]$$

where $\mathbf{I}$ is the identity matrix.

Proof: Since $\mathbf{U}_1$, $\mathbf{U}_2$ and $\mathbf{U}_3$ are not coplanar, then $(\mathbf{U}_1\times\mathbf{U}_2)$, $(\mathbf{U}_2\times\mathbf{U}_3)$ and $(\mathbf{U}_3\times\mathbf{U}_1)$ are not coplanar either since

$$(\mathbf{U}_1\times\mathbf{U}_2)\cdot[(\mathbf{U}_2\times\mathbf{U}_3)\times(\mathbf{U}_3\times\mathbf{U}_1)]$$

$$= (\mathbf{U}_1\times\mathbf{U}_2)\cdot[(\mathbf{U}_3\cdot(\mathbf{U}_1\times\mathbf{U}_2))\mathbf{U}_3] = [(\mathbf{U}_1\times\mathbf{U}_2)\cdot\mathbf{U}_3]^2 \neq 0 \qquad [3\text{-}18]$$

Then from [3-15] we must have

$$\mathbf{R} = [\,\mathbf{V}_1\ \mathbf{V}_2\ \mathbf{V}_3\,]\,[\mathbf{U}_1\ \mathbf{U}_2\ \mathbf{U}_3]^{-1} \qquad [3\text{-}19]$$

and from [3-16] we must have

$$\mathbf{R} = [\,\mathbf{V}_2\times\mathbf{V}_3\ \mathbf{V}_3\times\mathbf{V}_1\ \mathbf{V}_1\times\mathbf{V}_2\,][\,\mathbf{U}_2\times\mathbf{U}_3\ \mathbf{U}_3\times\mathbf{U}_1\ \mathbf{U}_1\times\mathbf{U}_2\,]^{-1} \qquad [3\text{-}20]$$

Using the transpose of [3-19] to multiply [3-20] we can get

$$\mathbf{R}^T\mathbf{R} = ([\mathbf{U}_1\ \mathbf{U}_2\ \mathbf{U}_3]^{-1})^T[\,\mathbf{V}_1\ \mathbf{V}_2\ \mathbf{V}_3\,]^T\cdot$$

$$[\,\mathbf{V}_2\times\mathbf{V}_3\ \mathbf{V}_3\times\mathbf{V}_1\ \mathbf{V}_1\times\mathbf{V}_2\,][\,\mathbf{U}_2\times\mathbf{U}_3\ \mathbf{U}_3\times\mathbf{U}_1\ \mathbf{U}_1\times\mathbf{U}_2\,]^{-1} \qquad [3\text{-}21]$$

By using the vector identity

$$\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C}) = \mathbf{B}\cdot(\mathbf{C}\times\mathbf{A}) = \mathbf{C}\cdot(\mathbf{A}\times\mathbf{B}) \qquad [3\text{-}22]$$

it is easy to show that

$$[\,\mathbf{V}_1\ \mathbf{V}_2\ \mathbf{V}_3\,]^T[\,\mathbf{V}_2\times\mathbf{V}_3\ \mathbf{V}_3\times\mathbf{V}_1\ \mathbf{V}_1\times\mathbf{V}_2\,] = (\mathbf{V}_1\cdot(\mathbf{V}_2\times\mathbf{V}_3))\,\mathbf{I} \qquad [3\text{-}23]$$

$$[\,\mathbf{U}_1\ \mathbf{U}_2\ \mathbf{U}_3\,]^T[\,\mathbf{U}_2\times\mathbf{U}_3\ \mathbf{U}_3\times\mathbf{U}_1\ \mathbf{U}_1\times\mathbf{U}_2\,] = (\mathbf{U}_1\cdot(\mathbf{U}_2\times\mathbf{U}_3))\,\mathbf{I} \qquad [3\text{-}24]$$

thus we have [3-17]. Q.E.D.

The above theorem tells us that if $\mathbf{R}$ satisfies [3-16] for given three noncoplanar vectors $\mathbf{U}_i$, i = 1,2,3, then $\mathbf{R}^T\mathbf{R}$ must be a scalar matrix. Note that if $\mathbf{V}_1\cdot(\mathbf{V}_2\times\mathbf{V}_3) = 0$, we'll get the trivial situation $\mathbf{R} = 0$. We shall always assume this not be the case. If

$$\mathbf{V}_1\cdot(\mathbf{V}_2\times\mathbf{V}_3) = \mathbf{U}_1\cdot(\mathbf{U}_2\times\mathbf{U}_3) \qquad [3\text{-}25]$$

then $\mathbf{R}$ is orthonormal. However from [3-18] we can not prove that $\mathbf{R}$ is orthonromal. But if $\mathbf{R}$ satisfies [3-18] for any vectors $\mathbf{U}_i$, i = 1,2,3, then we can prove that $\mathbf{R}$ is orthonormal. Note that

$$(\mathbf{V}_1\times\mathbf{V}_2)\times(\mathbf{V}_1\times\mathbf{V}_3) = [\mathbf{V}_1\cdot(\mathbf{V}_2\times\mathbf{V}_3)]\mathbf{V}_3 \qquad [3\text{-}26]$$

and

$$(\mathbf{U}_1\times\mathbf{U}_2)\times(\mathbf{U}_1\times\mathbf{U}_3) = [\mathbf{U}_1\cdot(\mathbf{U}_2\times\mathbf{U}_3)]\mathbf{U}_3 \quad [3\text{-}27]$$

by requiring

$$(\mathbf{V}_1\times\mathbf{V}_2)\times(\mathbf{V}_1\times\mathbf{V}_3) = \mathbf{R}\,[(\mathbf{U}_1\times\mathbf{U}_2)\times(\mathbf{U}_1\times\mathbf{U}_3)] \quad [3\text{-}28]$$

and comparing [3-15] we can get [3-25] and hence the orthonormality. Thus in general the orthonormality defined in Fact 2.1 is implied by the rigidity condition and hence redudent. And then if [3-26] is satisfied we shall get the Property 3 in Fact 2.1. That is equation [2-28]. Take the determinat at both sides of [3-20] we have

$$\det(\mathbf{R}) = \det([\,\mathbf{V}_1\;\mathbf{V}_2\;\mathbf{V}_3\,])\det([\mathbf{U}_1\;\mathbf{U}_2\;\mathbf{U}_3]^{-1})$$

$$= \frac{\mathbf{V}_1\cdot(\mathbf{V}_2\times\mathbf{V}_3)}{\mathbf{U}_1\cdot(\mathbf{U}_2\times\mathbf{U}_3)} = 1 \quad [3\text{-}29]$$

One can also show that from the orthonormality and the unit determinant property of $\mathbf{R}$ he can deduce equation [2-28], i.e., our definition of the rigidity. In stead of giving a mathematical proof here we adopt the well-known fact that [2-25] and [2-27] are the necessary and sufficeint conditions for a rigid motion. One can prove that [2-25] and [2-27] will imply [2-28] by following the procedures in Theorem 5.6. in section 5. Thus we have the following theorem:

Theorem 3.2.

The rigidity definition in Property 4 of Fact 2.1 is equivalent to the traditional rigidty definition [2-25] and [2-27]. ∎

So far we have got enough equations from the four correspondences in terms of the depths, one may immediately ask whether we can get more equations than unknowns in terms of motion parameters. One direct way is to represent the depths in terms of the motion parameters thus we may get 11 equations for 5 unknown motion parameters. However the equations got in this way may be too complicate to be used in practice. In the following we shall get some equations in terms of motion parameters only. But we cannot prove that all of them are independent or not.

Replacing [3-8] into [3-3] and [3-4] will lead to very complicate representations thus we shall only do the other transformation here. In practice if one really wants to do the replacement one'd better use [2-55] and [2-56]. Replacing $\mathbf{T} = Z'_1\mathbf{\Theta}'_1 - Z_1\mathbf{R}\mathbf{\Theta}_1$ into [3-2] for $i = 2$ we get

$$\Theta'_2 \cdot (Z'_1\Theta'_1 \times \mathbf{R}\Theta_2 - Z_1\mathbf{R}\Theta_1 \times \mathbf{R}\Theta_2) = 0 \qquad [3\text{-}30]$$

Using [2-28] and vector identity [3-22] we have

$$Z'_1\mathbf{R}\Theta_2 \cdot (\Theta'_2 \times \Theta'_1) = Z_1\Theta'_2 \cdot (\mathbf{R}(\Theta_1 \times \Theta_2)) \qquad [3\text{-}31]$$

Similarly if we replace $\mathbf{T} = Z'_1\Theta'_1 - Z_1\mathbf{R}\Theta_1$ into [3-2] for $i = 3$ we'll get

$$Z'_1\mathbf{R}\Theta_3 \cdot (\Theta'_3 \times \Theta'_1) = Z_1\Theta'_3 \cdot (\mathbf{R}(\Theta_1 \times \Theta_3)) \qquad [3\text{-}32]$$

The same way we can replace $\mathbf{T} = Z'_2\Theta'_2 - Z_2\mathbf{R}\Theta_2$ into [3-2] for $i = 1,3$ and $\mathbf{T} = Z'_3\Theta'_3 - Z_3\mathbf{R}\Theta_3$ into [3-2] for $i = 1,2$ to get the following general equation:

$$Z'_i\mathbf{R}\Theta_j \cdot (\Theta'_j \times \Theta'_i) = Z_i\Theta'_j \cdot (\mathbf{R}(\Theta_i \times \Theta_j)),\ j \neq i,\ i,j = 1,2,3 \qquad [3\text{-}33]$$

Obviously [3-13] gives 6 equations for the depths if we represent $\mathbf{R}$ in the form of [3-6]. And besides, if we delete Z'_1 and Z_1 from [3-11] and [3-12] and do the same things for other equations in [3-33] we'll get

$$(\mathbf{R}\Theta_j \cdot (\Theta'_j \times \Theta'_i))\,(\Theta'_k \cdot (\mathbf{R}(\Theta_i \times \Theta_k))) = (\mathbf{R}\Theta_k \cdot (\Theta'_k \times \Theta'_i))\,(\Theta'_j \cdot (\mathbf{R}(\Theta_i \times \Theta_j)))$$
$$i \neq j \neq k,\ i,j,k = 1,2,3 \qquad [3\text{-}34]$$

[3-34] gives 3 equations for the rotation matrix! But these equations turns out to be dependent. To prove this let's rearrange [3-2] as following

$$\mathbf{T}^T(\Theta'_i \times \mathbf{R}\Theta_i) = 0,\ i = 1,2,\cdots,n \qquad [3\text{-}35]$$

Given three correspondences we must have

$$\mathbf{T}^T[\,\Theta'_1 \times \mathbf{R}\Theta_1\ \ \Theta'_2 \times \mathbf{R}\Theta_2\ \ \Theta'_3 \times \mathbf{R}\Theta_3\,] = 0 \qquad [3\text{-}36]$$

From linear algebra we know that to make $\mathbf{T}$ have a nonzero unique solution up to a scalar we must have

$$\mathrm{rank}\,([\,\Theta'_1 \times \mathbf{R}\Theta_1\ \ \Theta'_2 \times \mathbf{R}\Theta_2\ \ \Theta'_3 \times \mathbf{R}\Theta_3\,]) = 2 \qquad [3\text{-}37]$$

or

$$(\Theta'_1 \times \mathbf{R}\Theta_1) \cdot [(\Theta'_2 \times \mathbf{R}\Theta_2) \times (\Theta'_3 \times \mathbf{R}\Theta_3)] = 0 \qquad [3\text{-}38]$$

After some manipulations one can find [3-38] is equivalent to [3-34]. So we have proved there is only one independent equation in [3-34]. Given three correspondences we also have another equation containing $\mathbf{R}$ as unknowns only. First we intrdouce the following vector identity:

$$(X'_i \times X'_j) \cdot (X'_i - X'_j) = 0, \quad \text{for any } i,j \qquad [3\text{-}39]$$

[3-39] is of no meaning when $i = j$. So in the following when [3-39] is used we always assume $i \neq j$.

In terms of Θ and using [3-1] we get

$$(\Theta'_i \times \Theta'_j)^T \mathbf{R} (\Theta_i Z_i - \Theta_j Z_j) = 0 \qquad [3\text{-}40]$$

Let

$$\beta_{ij} = \Theta'_i \times \Theta'_j \ , \quad \text{for any } i \neq j \qquad [3\text{-}41]$$

then, [3-40] becomes

$$\beta_{ij}^T \mathbf{R} \Theta_i Z_i = \beta_{ij}^T \mathbf{R} \Theta_j Z_j \qquad [3\text{-}42]$$

Considering i,j = 1,2,3 we may have

$$Z_1 \beta_{12}^T \mathbf{R} \Theta_1 = Z_2 \beta_{12}^T \mathbf{R} \Theta_2 \qquad [3\text{-}43]$$

$$Z_2 \beta_{23}^T \mathbf{R} \Theta_2 = Z_3 \beta_{23}^T \mathbf{R} \Theta_3 \qquad [3\text{-}44]$$

$$Z_3 \beta_{31}^T \mathbf{R} \Theta_3 = Z_1 \beta_{31}^T \mathbf{R} \Theta_1 \qquad [3\text{-}45]$$

Multiply [3-43], [3-44] and [3-45] we get

$$\beta_{12}^T \mathbf{R} \Theta_1 \cdot \beta_{23}^T \mathbf{R} \Theta_2 \cdot \beta_{34}^T \mathbf{R} \Theta_3 = \beta_{12}^T \mathbf{R} \Theta_2 \cdot \beta_{23}^T \mathbf{R} \Theta_3 \cdot \beta_{31}^T \mathbf{R} \Theta_1 \qquad [3\text{-}46]$$

[3-46] seems to be independent of [3-34]. Now let's consider the four point problem.

Immediately we see that if one more correspondence is introduced then we can get at least one more equation involving **R** as unknown only. Now if one more correspondence is introduced, then we'll get 2 more equations in the form [3-46] and at least one more equations in the form [3-34] with only 2 more unknown depths introduced. Similar to [3-37] by requiring

$$\text{rank}\left([\ \Theta'_1 \times \mathbf{R}\Theta_1 \ \ \Theta'_2 \times \mathbf{R}\Theta_2 \ \ \Theta'_3 \times \mathbf{R}\Theta_3 \ \ \Theta'_4 \times \mathbf{R}\Theta_4\ ] \right) = 2 \qquad [3\text{-}47]$$

we'll get up to two independent equations in the form of

$$(\mathbf{R}\Theta_j \cdot (\Theta'_j \times \Theta'_i))\,(\Theta'_k \cdot (\mathbf{R}(\Theta_i \times \Theta_k))) = (\mathbf{R}\Theta_k \cdot (\Theta'_k \times \Theta'_i))\,(\Theta'_j \cdot (\mathbf{R}(\Theta_i \times \Theta_j)))$$

$$i \neq j \neq k,\ \ i,j,k = 1,2,3,4 \qquad [3\text{-}48]$$

Also similar to [3-46] we have following equations containing **R** only

$$\beta_{23}^T \mathbf{R} \Theta_2 \cdot \beta_{34}^T \mathbf{R} \Theta_3 \cdot \beta_{42}^T \mathbf{R} \Theta_4 = \beta_{23}^T \mathbf{R} \Theta_3 \cdot \beta_{34}^T \mathbf{R} \Theta_4 \cdot \beta_{42}^T \mathbf{R} \Theta_2 \qquad [3\text{-}49]$$

$$\beta_{21}^T \mathbf{R} \Theta_2 \cdot \beta_{14}^T \mathbf{R} \Theta_1 \cdot \beta_{42}^T \mathbf{R} \Theta_4 = \beta_{21}^T \mathbf{R} \Theta_1 \cdot \beta_{14}^T \mathbf{R} \Theta_4 \cdot \beta_{42}^T \mathbf{R} \Theta_2 \qquad [3\text{-}50]$$

$$\beta_{13}^T \mathbf{R} \Theta_1 \cdot \beta_{34}^T \mathbf{R} \Theta_3 \cdot \beta_{41}^T \mathbf{R} \Theta_4 = \beta_{13}^T \mathbf{R} \Theta_3 \cdot \beta_{34}^T \mathbf{R} \Theta_4 \cdot \beta_{41}^T \mathbf{R} \Theta_1 \qquad [3\text{-}51]$$

The reason we do not include

$$\beta_{23}^T \mathbf{R} \Theta_2 \cdot \beta_{31}^T \mathbf{R} \Theta_3 \cdot \beta_{12}^T \mathbf{R} \Theta_1 = \beta_{23}^T \mathbf{R} \Theta_3 \cdot \beta_{31}^T \mathbf{R} \Theta_1 \cdot \beta_{12}^T \mathbf{R} \Theta_2 \qquad [3\text{-}52]$$

is because [3-52] can be deduced from [3-49] ~ [3-51]. However we cannot prove that the equations in [3-49] ~ [3-51] are independent from those two in [3-48].

Note that the labeling in equation [3-19] ~ [3-24] is arbitrary, thus we can state the equations in [3-49] to [3-51] in a general form. The equations

$$\beta_{ij}^T \mathbf{R}\Theta_i \cdot \beta_{jk}^T \mathbf{R}\Theta_j \cdot \beta_{ki}^T \mathbf{R}\Theta_k = \beta_{ij}^T \mathbf{R}\Theta_j \cdot \beta_{jk}^T \mathbf{R}\Theta_k \cdot \beta_{ki}^T \mathbf{R}\Theta_i$$

$$i \neq j \neq k, \ \ i,j,k = 1,2,3,4 \qquad [3\text{-}53]$$

will give up to three independent equations containing **R** as unknowns only.

So far we have got many equations containing depths or motion parameters only. All the equations listed here are just typical. One may find out many other equivalent equations of other forms. If it turns out that we have more independent equations than unknows then a finite number of solutions may be got from numerical methods. Once a finite number of solutions have been achieved, many other constraints can be adopted to figure out the spurious solutions if there is any. For example, the positive depth constraint

$$Z_i > 0, \ \ Z'_i > 0, \ \ i = 1,2,3 \qquad [3\text{-}54]$$

and the small motion constraint

$$(\Theta'_i Z'_i - \Theta'_j Z'_j) \cdot (\Theta_i Z_i - \Theta_j Z_j) > 0, \ \ i \neq j, \ \ i,j = 1,2,3,4 \qquad [3\text{-}55]$$

The nonlinear equations seem more feasible if one use equation [3-48] and [3-54] since it only contains the rotation matrix as unknowns. And the dimension should not be too high to get reliable numerical results if one changes it into the polynomial form. Many things are left undone here. Because of the high dimension a closed form solution is certainly not availabe. So in general only numerical solutions can be got.

Even if we have found some way to solve the nonlinear equations in terms of depths or motion parameters for just four correspondences, another big problem is how to make the solution robust if we have many other correspondences. And what performance criterion should be used in the nonlinear equation solutions? Can we search some reasonable space to get a globally optimal solution from the nonlinear approach for any performance criterion? We impose these questions to arise other researchers' interests. If we cannot solve these problem then it seems our algorithm introduced in section 8 could be a good approach.

From these nonlinear equations we can analyse many special surface configurations, such as planar correspondences, colinear correspondences with some correspondences aside, etc. However this analysis will be beyond the scope of this paper.

4. Approximation Approach For Plane Motion

In this section we introduce a closed form solution of plane motion with small rotation angles. This solution, though not perfect, is a quite direct and fast solution of the motion. We have found some usage for it in some other places, but we only introduce the algorithm itself here. We find it works well for small rotation.

Assume the plane has an equation

$$\begin{aligned} Z &= p\,X + q\,Y + Z_0 \\ &= p\,x\,Z + q\,y\,Z + Z_0 \end{aligned} \qquad [4\text{-}1]$$

with $Z_0 \neq 0$, or

$$\frac{Z_0}{Z} = 1 - p\,x - qy \qquad [4\text{-}2]$$

let

$$u_1 = -p\,\frac{\dot{X}_c}{Z_0} - \frac{\dot{Z}_c}{Z_0} \stackrel{\Delta}{=} -p\,V_X - V_Z \qquad [4\text{-}3]$$

$$u_2 = -q\,\frac{\dot{X}_c}{Z_0} - \omega_Z \stackrel{\Delta}{=} -q\,V_X - \omega_Z \qquad [4\text{-}4]$$

$$u_3 = \frac{\dot{X}_c}{Z_0} + \omega_Y \stackrel{\Delta}{=} V_X + \omega_Y \qquad [4\text{-}5]$$

$$u_4 = -p\,\frac{\dot{Y}_c}{Z_0} + \omega_Z \stackrel{\Delta}{=} -p\,V_Y + \omega_Z \qquad [4\text{-}6]$$

$$u_5 = -q\,\frac{\dot{Y}_c}{Z_0} - \frac{\dot{Z}_c}{Z_0} \stackrel{\Delta}{=} -q\,V_Y - V_Z \qquad [4\text{-}7]$$

$$u_6 = \frac{\dot{Y}_c}{Z_0} - \omega_X \stackrel{\Delta}{=} V_Y - \omega_X \qquad [4\text{-}8]$$

$$u_7 = -p\,\frac{\dot{Z}_c}{Z_0} - \omega_Y \overset{\Delta}{=} -p\,V_Z - \omega_Y \tag{4-9}$$

$$u_8 = -q\,\frac{\dot{Z}_c}{Z_0} + \omega_X \overset{\Delta}{=} -q\,V_Z + \omega_X \tag{4-10}$$

and

$$\mathbf{U} = (\,u_1\ u_2\ u_3\ u_4\ u_5\ u_6\ u_7\ u_8\,)^T \tag{4-11}$$

then [2-17] becomes

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} x & y & 1 & 0 & 0 & 0 & -x^2 & -xy \\ 0 & 0 & 0 & x & y & 1 & -xy & -y^2 \end{bmatrix} \mathbf{U} \tag{4-12}$$

$$\begin{aligned} x' &= x + \dot{x}\Delta t, & \dot{x} &= (x' - x)/\Delta t \\ y' &= y + \dot{y}\Delta t, & \dot{y} &= (y' - y)/\Delta t \end{aligned} \tag{4-13}$$

Given four point correspondences $(x_i, y_i),(x_i', y_i')$, $i = \{1,2,3,4\}$, we can solve U by

$$\mathbf{U} = \mathbf{H}_1^{-1} \begin{bmatrix} \dot{x}_1 \\ \dot{y}_1 \\ \cdot \\ \cdot \\ \dot{x}_4 \\ \dot{y}_4 \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1^2 & -x_1y_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -x_1y_1 & -y_1^2 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & & \cdot & & & \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -x_4^2 & -x_4y_4 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -x_4y_4 & -y_4^2 \end{bmatrix}^{-1} \begin{bmatrix} \dot{x}_1 \\ \dot{y}_1 \\ \cdot \\ \cdot \\ \dot{x}_4 \\ \dot{y}_4 \end{bmatrix} \tag{4-14}$$

A direct rank checking procedure shows that iff no three points of (x_i,y_i), $i = 1,2,3,4$, are colinear in the image plane then the inversion of $\mathbf{H}_1$ exists.

After we get U we can recover the motion and structure parameters from U to get up to two sets of solutions. From [2-13], we can write

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} 1-pV_X & -\omega_z-qV_X & \omega_Y+V_X \\ +\omega_Z-pV_Y & 1-qV_Y & V_Y-\omega_X \\ -\omega_Y-pV_Z & \omega_X-qV_Z & 1+V_Z \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \mathbf{K} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} k_1 & k_2 & k_3 \\ k_4 & k_5 & k_6 \\ k_7 & k_8 & k_9 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{4-15}$$

It turns out

$$k_2 = u_2;\ \ k_3 = u_3;\ \ k_4 = u_4;$$

$$k_6 = u_6;\ \ k_7 = u_7;\ \ k_8 = u_8;$$

$$k_1 = 1-pV_X;\ \ k_5 = 1-qV_Y;\ \ k_9 = 1+V_Z; \qquad [4\text{-}16]$$

$$k_1 - k_9 = u_1;\ \ k_5 - k_9 = u_5;$$

For our purpose, we rewrite [4-15] as

$$Z' \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = K \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} Z \qquad [4\text{-}17]$$

or

$$x' = \frac{k_1 x + k_2 y + k_3}{k_7 x + k_8 y + k_9}$$

$$y' = \frac{k_4 x + k_5 y + k_6}{k_7 x + k_8 y + k_9} \qquad [4\text{-}18]$$

Given four correspondences (x_i, y_i), (x_i', y_i'), $i = \{1,2,3,4\}$, we can solve K to within a scalar factor from

$$\mathbf{H}_2 \begin{bmatrix} k_1 \\ k_2 \\ \cdot \\ \cdot \\ \cdot \\ k_9 \end{bmatrix} \overset{\Delta}{=} \begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1x_1' & -y_1x_1' & -x_1' \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -x_1y_1' & -y_1y_1' & -y_1' \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & & & \cdot & & & \\ & & & & & \cdot & & & \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -x_4x_4' & -y_4x_4' & -x_4' \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -x_4y_4' & -y_4y_4' & -y_4' \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ \cdot \\ \cdot \\ \cdot \\ k_9 \end{bmatrix} = 0 \qquad [4\text{-}19]$$

A direct rank checking procedure shows that the rank of $\mathbf{H}_2$ has a rank 8 iff no three points in either image are colinear (see also [23]). If the projection of the plane is not a line in the image plane, that is, the projection is not *degenerate* , then, the colinearity of three points in the image is equivalent to that of their original duals on the plane in space. So if the projection is not degenerate, then iff no three points are colinear in the space, $\mathbf{H}_2$ will have full row rank. We shall always assume this condition be satisfied in the plane model discussion. Since the projection of a line in the space must be a line in the image, thus a sufficient condition to guarantee H_1 and H_2 have rank 8 is to make sure no three points are colinear in the image before and after motion.

Given four correspondences (x_i, y_i), (x_i', y_i'), $i = \{1,2,3,4\}$, we can uniquely solve U from [4-14] and K to within a scalar factor from [4-19]. Let one of the solution of [4-19] be k_i''s, $i = 1,\cdots,9$, then

$$k_i' = \lambda k_i, \ i=1,\cdots,9 \qquad [4\text{-}20]$$

and suppose one of $u_2, u_3, u_4, u_6, u_7, u_8$, say u_2 , not be zero, then from [4-16] we know V_Z can be uniquely decided by

$$V_Z = k_9' u_2 / k_2' - 1; \qquad [4\text{-}21]$$

However, [4-21] may not give accurate results since U calculated from [4-12] is too sensitive to correspondence errors. The reason is that [4-12] comes from [2-17] which is again an approximation formula. So we'd better use k_i''s only to solve motion parameters. Hence we adopt the following algorithm.

Let

$$u_1' = k_1' - k_9'; \ u_2' = k_2'; \ u_3' = k_3';$$
$$u_4' = k_4'; \ u_5' = k_5' - k_9'; \ u_6' = k_6'; \qquad [4\text{-}22]$$
$$u_7' = k_7'; \ u_8' = k_8';$$

Obviously

$$u_i' = \lambda u_i, \ i=1,\cdots,8, \text{ for some } \lambda$$

Let

$$s_1 \triangleq u_2' + u_4' = -pV_Y' - qV_X' \qquad [4\text{-}23]$$

$$s_2 \triangleq u_8' + u_6' = -qV_Z' + V_Y' \qquad [4\text{-}24]$$

$$s_3 \triangleq u_7' + u_3' = -pV_Z' + V_X' \qquad [4\text{-}25]$$

with

$$V_X' = \lambda V_X; \ V_Y' = \lambda V_Y; \ V_Z' = \lambda V_Z; \qquad [4\text{-}26]$$

and p, q being unaffected. Then we have

$$V_X' = pV_Z' + s_3 \qquad [4\text{-}27]$$

$$V_Y' = qV_Z' + s_2 \qquad [4\text{-}28]$$

Substitute [4-26] and [4-27] into [4-22], [4-3] and [4-8], we get

$$p\,(\,qV_Z'+s_2\,)+q\,(\,pV_Z'+s_3\,)\;=\;-s_1 \tag{4-29}$$

$$p\,(\,pV_Z'+s_3\,)+V_Z'\;=\;-u_1' \tag{4-30}$$

$$q\,(\,qV_Z'+s_2\,)+V_Z'\;=\;-u_5' \tag{4-31}$$

Summing [4-24], [4-25] and [4-26] leads to

$$V_Z'\,(p+q)^2+(s_2+s_3)\,(p+q)+2V_Z'+s_1+u_1'+u_5'=0 \tag{4-32}$$

From [4-30] ~ [32], we shall have

$$p\;=\;(-s_3\pm t_1)\,/\,2V_Z' \tag{4-33}$$

$$q\;=\;(-s_2\pm t_2)\,/\,2V_Z' \tag{4-34}$$

$$p+q\;=\;(-(s_2+s_3)\pm t_3)\,/\,2V_Z' \tag{4-35}$$

where

$$t_1\;=\;\sqrt{s_3^2-4V_Z'(V_Z'+u_1')} \tag{4-36}$$

$$t_2\;=\;\sqrt{s_2^2-4V_Z'(V_Z'+u_5')} \tag{4-37}$$

$$t_3\;=\;\sqrt{(s_2+s_3)^2-4V_Z'(2V_Z'+s_1+u_1'+u_5')} \tag{4-38}$$

Thus, we must have

$$\pm\,t_1\pm t_2\;=\;\pm t_3 \tag{4-39}$$

Squaring [4-39] twice leads to

$$V_Z'[4V_Z'^{\,3}+4V_Z'^{\,2}(u_1'+u_5')-V_Z'(s_1^2+s_2^2+s_3^2-4u_1'u_5')-u_1's_2^2-u_5's_3^2+s_1s_2s_3]=0 \tag{4-40}$$

There are two different situations:

1. $V_Z'=0$;

The condition for $V_Z'=0$, from [4-30] ~ [4-32], is

$$u_1's_2^2+u_5's_3^2\;=\;s_1s_2s_3 \tag{4-41}$$

In this case, p and q can be uniquely decided from [5-11] and [5-12] by

$$p\;=\;-\,u_1'\,/\,s_3 \tag{4-42}$$

$$q\;=\;-\,u_5'\,/\,s_2 \tag{4-43}$$

and from [4-16] we know

$$\lambda\;=\;k_9' \tag{4-44}$$

Then the motion parameters can be got from [4-3] ~ [4-10], for example,

$$V_Y = s_2 / \lambda$$
$$V_X = s_3 / \lambda$$
$$\omega_X = u_8' / \lambda \qquad [4\text{-}45]$$
$$\omega_Y = -u_7' / \lambda$$
$$\omega_Z = -u_2' / \lambda - qV_X$$

This result suggests that if we want to decide the plane structure and motion parameters uniquely from two views the motion should not include a translation along optical axis. This suggests an application in stereo vision. We are currently carrying on research in this direction.

2. $V_Z' \neq 0$;

Then V_Z' must be a root of

$$4V_Z'^3 + 4V_Z'^2(u_1'+u_5') - V_Z'(s_1^2 + s_2^2 + s_3^2 - 4u_1'u_5') - u_1's_2^2 - u_5's_3^2 + s_1s_2s_3] = 0 \qquad [4\text{-}46]$$

However there are up to 3 real roots of [4-46]. We can use following constraints to reduce up to two fake roots:

$$s_3^2 - 4V_Z'(V_Z'+u_1') \geq 0$$
$$s_2^2 - 4V_Z'(V_Z'+u_5') \geq 0 \qquad [4\text{-}47]$$
$$(s_2+s_3)^2 - 4V_Z'(2V_Z'+s_1+u_2'+u_6') \geq 0$$

But we still cannot assure V_Z' be unique. So we have to use the following method to find the right V_Z'. After we get V_Z' we can solve λ from

$$\lambda = k_9' - V_Z' \qquad [4\text{-}48]$$

Assume u_i, $i = \{ 1, \dots, 9 \}$, be calculated from [4-14] with the given four correspondences. Then the right V_Z' and hence the correct λ should make

$$\sum_{i=1}^{8} (\lambda - \frac{u_i'}{u_i})^2 = \text{minimum} \qquad [4\text{-}49]$$

This way we decide the correct V_Z'. After we get the right V_Z' we can now solve other parameters, though there are two solutions of them.

A. if

$$t_1 + t_2 = t_3 \qquad [4\text{-}50]$$

then from [4-33] to [4-35] we know

$$p_{1,2} = (-s_3 \pm t_1) / 2V_Z' \qquad [4\text{-}51]$$

$$q_{1,2} = (-s_2 \pm t_2) / 2V_Z' \qquad [4\text{-}52]$$

B. otherwise there must be

$$t_1 - t_2 = t_3 \text{, or } t_1 - t_2 = -t_3 \qquad [4\text{-}53]$$

then

$$p_{1,2} = (-s_3 \pm t_1) / 2V_Z' \qquad [4\text{-}54]$$

$$q_{1,2} = (-s_2 \mp t_2) / 2V_Z' \qquad [4\text{-}55]$$

After p and q are decided the motion parameters will be given, for example, by

$$(V_X)_{1,2} = (p_{1,2}V_Z' + s_3) / \lambda$$

$$(V_Y)_{1,2} = (q_{1,2}V_Z' + s_2) / \lambda$$

$$(\omega_X)_{1,2} = -u_6' / \lambda + (V_Y)_{1,2} \qquad [4\text{-}56]$$

$$(\omega_Y)_{1,2} = -(V_X)_{1,2} + u_3' / \lambda$$

$$(\omega_Z)_{1,2} = -u_2' / \lambda - q_{1,2}(V_X)_{1,2}$$

The spurious solution can be figured out by requiring ω_X, ω_Y, ω_Z be small. Or if multiple plane patches of the same object are considered at the same time, we can use the motion consistence condition and surface continuity condition to find the right solution. We'll discuss this problem later.

Simulation results show that this algorithm works well only for small rotation. And in general cases it will only guarantee the accuracy of the dominant motion parameters. It is less robust than the algorithm discussed in the next section. The experiment results using this model are given in Table ?. From our experience, the velocity decomposition model generally cannot guarantee the accuracy of all the motion parameters, especially when the rotation is large. So if one insists to get the velocities $\dot{X}_c$, $\dot{Y}_c$, and ω_X etc., instead of **R** and **T**, we still suggest that one solve **R** and **T** first, and then use Fact 2.4 to get the velocities.

5. Perfect Transformation Approach For Plane Motion

In this section we introduce a robust algorithm for the plane model solution for both motion representations. Some of the theorems we shall give here were established by Tsai ([18],1982) and Weng ([25],1988) earlier with very complicated proofs. We repeat them here by alternative but much briefer and stronger proofs to make our other theorems and our algorithm better understood . We'll give the references for those theorems established by other people in a different way. The algorithm introduced here gives robust solutions for plane motion. Thus it may be directly useful in many applications where plane models apply. For example, in industry robot, aircraft and automobile navigation etc., we may sometimes directly use the plane model to solve motion. However if we do not have any information about the surface we cannot use it to directly solve general surface motion. But the goal of our approach is to use this model to do the job of general surface motion solution. So we'll give more than enough results for plane motion solution.

Rewrite [4-1] as

$$N^T \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \triangleq \begin{bmatrix} n_1 \, n_2 \, n_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} -\frac{p}{Z_0} & -\frac{q}{Z_0} & \frac{1}{Z_0} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = 1 \qquad [5\text{-}1]$$

then [2-18] becomes

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = R \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + T N^T \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = K \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \qquad [5\text{-}2]$$

where

$$K = R + T N^T = \begin{bmatrix} k_1 & k_2 & k_3 \\ k_4 & k_5 & k_6 \\ k_7 & k_8 & k_9 \end{bmatrix} \qquad [5\text{-}3]$$

We see [5-2] and [5-3] are in the same form as [4-15] though the inner representation is different. However the value of both **K**'s should be the same. Thus we will consider **K** is the same. As discussed earlier, **K** can be solved to within a constant with four correspondences from [4-19]. The scalar can be decided by the rigidity condition (Tsai,[18] and Weng [25]).

Before we go into detail to solve **R**, **T**, and **N** from **K** we should note the duality of the problem. Let's consider

$$\mathbf{K}^T = \mathbf{R}^T + \mathbf{N}\mathbf{T}^T \qquad [5\text{-}4]$$

we immediately see that to recover $\mathbf{R}$, $\mathbf{T}$, and $\mathbf{N}$ from $\mathbf{K}$ is equivalent to recover $\mathbf{R}^T$, $\mathbf{N}$, and $\mathbf{T}$ from $\mathbf{K}^T$. Because $\mathbf{K}\mathbf{K}^T$ has exactly the same eigenvalues as $\mathbf{K}^T\mathbf{K}$, we can see the discussion below about the property of $\mathbf{N}$ also suits to $\mathbf{T}$. We will call this property as the *duality property*. One difference should be remembered is that $\mathbf{T}$ can be anything, but $\mathbf{N}$ can never be zero.

Now we show how to decide $\mathbf{K}$ uniquely. We first give following theorem.

Theorem 5.1. (See also Weng, [25])

The necessary and sufficient condition for a 3-dimensional matrix $\mathbf{K}$ to be able to decomposed into the form [5-3] with $\mathbf{R}$ a rotation matrix, $\mathbf{T}$ and $\mathbf{N}$ vectors, is that the three eigenvalues of the matrix $\mathbf{K}^T\mathbf{K}$ λ_1, λ_2, and λ_3 with $\lambda_1 \le \lambda_2 \le \lambda_3$ satisfy

$$0 \le \lambda_1 \le \lambda_2 = 1 \le \lambda_3 \qquad [5\text{-}5]$$

Proof: We only prove the necessary part here, the sufficient part is automatically proved by our following constructing procedure. Since the eigenvalues of $\mathbf{K}^T\mathbf{K}$ must be nonnegative, what we need to do is to show $\lambda_2 = 1$. First we show 1 is an eigenvalue of $\mathbf{K}^T\mathbf{K}$. In any case, let any nonzero real vector $\mathbf{X}$ satisfy:

$$\mathbf{N}^T\mathbf{X} = 0, \text{ and } \mathbf{T}^T\mathbf{R}\,\mathbf{X} = 0 \qquad [5\text{-}6]$$

then

$$\begin{aligned} \mathbf{K}^T\mathbf{K}\,\mathbf{X} &= (\mathbf{R}^T + \mathbf{N}\mathbf{T}^T)(\mathbf{R} + \mathbf{T}\mathbf{N}^T)\mathbf{X} \\ &= (\mathbf{R}^T + \mathbf{N}\mathbf{T}^T)\mathbf{R}\mathbf{X} = \mathbf{R}^T\mathbf{R}\mathbf{X} = \mathbf{X} = 1\cdot\mathbf{X} \end{aligned} \qquad [5\text{-}7]$$

hence $\lambda = 1$ is an eigenvalue. Let it be λ_2. Then we show one of the rest two eigenvalues, say, λ_1 must be less than or equal to 1 and the other, say, λ_3 must be larger than or equal to 1. Let $\mathbf{W}$ be an orthonormal matrix such that

$$\mathbf{W}^T\mathbf{K}^T\mathbf{K}\,\mathbf{W} = diag\,(\lambda_1\,\lambda_2\,\lambda_3), \qquad [5\text{-}8]$$

let $\mathbf{X}$ be any non-zero vector orthogonal to $\mathbf{N}$, i.e.,

$$\mathbf{N}^T\mathbf{X} = 0 \qquad [5\text{-}9]$$

then

$$\mathbf{K}\mathbf{X} = \mathbf{R}\mathbf{X} \qquad [5\text{-}10]$$

and let $\mathbf{U} = [\,u_1\,u_2\,u_3\,]^T = \mathbf{W}^T\mathbf{X}$, hence $\mathbf{X} = \mathbf{W}\mathbf{U}$, thus from [5-10] and the orthonormality condition [2-25] we should have

$$\mathbf{X}^T \mathbf{K}^T\mathbf{K}\mathbf{X} = \mathbf{X}^T\mathbf{R}^T\mathbf{R}\mathbf{X} = \mathbf{X}^T\mathbf{X} \qquad [5\text{-}11]$$

or

$$\mathbf{U}^T \mathbf{W}^T\mathbf{W}\,\mathbf{U} = \mathbf{U}^T \mathbf{W}^T\mathbf{K}^T\mathbf{K}\mathbf{W}\,\mathbf{U} \qquad [5\text{-}12]$$

From [5-8] and the orthonormality of $\mathbf{W}$ we have

$$u_1^2 + u_2^2 + u_3^2 = \lambda_1 u_1^2 + \lambda_2 u_2^2 + \lambda_3 u_3^2 \qquad [5\text{-}13]$$

or

$$(1 - \lambda_1)\, u_1^2 = (\lambda_3 - 1) u_3^2 \qquad [5\text{-}14]$$

Because λ_1 and λ_3 are both nonnegative, to make equation [5-14] to be true, the larger one, say λ_3 must be greater than or equal to 1 and the smaller one, say λ_1 must be less than or equal to 1. Q.E.D..

Immediately we have the following Corollaries.

Corollary 5.1 .

The other two eigenvalues are given by

$$\lambda_1 = \frac{b - \sqrt{b^2-4c}}{2}, \quad \lambda_3 = \frac{b + \sqrt{b^2-4c}}{2} \qquad [5\text{-}15]$$

where

$$b = Trace(\mathbf{K}^T\mathbf{K}) - 1, \text{ and } c = \det(\mathbf{K}^T\mathbf{K}) \stackrel{\Delta}{=} |\mathbf{K}|^2 \qquad [5\text{-}16]$$

Proof: From linear algebra we know

$$\lambda_1 \lambda_2 \lambda_3 = |\mathbf{K}^T\mathbf{K}| = |\mathbf{K}|^2 \text{ and } \lambda_1 + \lambda_2 + \lambda_3 = Trace(\mathbf{K}^T\mathbf{K}) \qquad [5\text{-}17]$$

Since $\lambda_2 = 1$ thus we have [5-15]. Q.E.D..

Corollary 5.2 .

The trace and determinant of $\mathbf{K}^T\mathbf{K}$ must satisfy

$$(Trace(\mathbf{K}^T\mathbf{K})-1)^2 \geq 4|\mathbf{K}|^2 \qquad [5\text{-}18]$$

Proof : this is because both λ_1 and λ_3 in [5-15] must be real numbers. Q.E.D..

Corollary 5.3 .

In any case the three eigenvalues of $\mathbf{K}^T\mathbf{K}$ are equal, they must be 1 and in this situation $|\mathbf{K}|=1$, $Trace(\mathbf{K}^T\mathbf{K}) = 3$ and if two of them are equal they must be 1 either.

Proof: This is a direct consequence of Theorem 5.1 and [5-15]. Q.E.D..

Corollary 5.4 .

The rank of $\mathbf{K}$ is at least 2.

Proof: Since $\mathbf{K}^T\mathbf{K}$ has at most one eigenvalue less than 1 thus it can at most has one zero eigenvalue. Hence its rank is at least 2. Because $\mathbf{K}$ has the same rank as $\mathbf{K}^T\mathbf{K}$ thus the rank of $\mathbf{K}$ is at least 2. Q.E.D..

It seems [5-15] gives a direct way to find the eigenvalues of $\mathbf{K}^T\mathbf{K}$. However we can hardly use it, because from [4-19] we can only solve $\mathbf{K}$ up to a scalar, that is we can only get $\mathbf{K}' = \alpha\mathbf{K}$. So what we can do is to first solve all the eigenvalues $\lambda'_1, \lambda'_2, \lambda'_3$ of $\mathbf{K}'^T\mathbf{K}'$, assuming $\lambda'_1 \le \lambda'_2 \le \lambda'_3$ and then to get

$$\mathbf{K} = \frac{\mathbf{K}'}{\pm\sqrt{\lambda'_2}} \qquad [5\text{-}19]$$

To decide whether $+\sqrt{\lambda'_2}$ or $-\sqrt{\lambda'_2}$ is to be used in [5-19], we need to make use of one point correspondence in the image. Let (x, y) be such a point. Because it's depth Z and its correspondence point's depth Z′ must be positive we shall have

$$\frac{Z'}{Z} = k_7x + k_8y + k_9 > 0 \qquad [5\text{-}20]$$

So in [5-19] one must select $+\sqrt{\lambda'_2}$ or $-\sqrt{\lambda'_2}$ such that [5-20] is satisfied. Then we can get

$$\lambda_1 = \lambda'_1 / \lambda'_2, \quad \lambda_2 = 1, \quad \lambda_3 = \lambda'_3 / \lambda'_2 \qquad [5\text{-}21]$$

Though $\mathbf{K}$ can be solved uniquely, however, we may have up to two sets of solutions to recover $\mathbf{R}$, $\mathbf{T}$, and $\mathbf{N}$ from $\mathbf{K}$. The spurious solution can only be figured out by other constraints. Tsai ([18]) and Weng ([25]) worked out an algorithm separately. Based on our theorems we suggest a robust algorithm here. This algorithm is a quite direct solution.

Theorem 5.2 . (see also Tsai, [18] and Weng, [25])

For any matrix $\mathbf{K}$ of the form [5-3],

1. If two of the eigenvalues of $\mathbf{K}^T\mathbf{K}$ are equal to 1 then $\mathbf{N}$ is

uniquely decided up to a scalar.

2. If only one of the eigenvalues of $\mathbf{K}^T\mathbf{K}$ are equal to 1 then there are two independent sets of solution for $\mathbf{N}$, each up to a scalar.

3. If all three eigenvalues of $\mathbf{K}^T\mathbf{K}$ are equal to 1 then there are an infinite number of solutions for $\mathbf{N}$, that is, $\mathbf{N}$ is undecided.

Proof: 1. Look at the equation of [5-14]. When either λ_1 or λ_3 is 1 then [5-14] defines one plane in terms of $\mathbf{U}$. Neglecting the sign we can write [5-14] as

$$[\ \sqrt{1-\lambda_1}\ \ 0\ \ \sqrt{\lambda_3-1}\ ]\mathbf{U} = 0 \qquad [5\text{-}22]$$

remember that $\mathbf{U} = \mathbf{W}^T\mathbf{X}$, thus [5-18] becomes

$$[\ \sqrt{1-\lambda_1}\ \ 0\ \ \sqrt{\lambda_3-1}\ ]\mathbf{W}^T\mathbf{X} = 0 \qquad [5\text{-}23]$$

By comparing the requirement [5-9] with [5-23] we know that the two plane must be the same except a scalar since $\mathbf{X}$ is chosen as any point on the plane defined by [5-9]. Thus $\mathbf{N}$ is decided up to a scalar by

$$\mathbf{N}^T = [\ \sqrt{1-\lambda_1}\ \ 0\ \ \sqrt{\lambda_3-1}\ ]\mathbf{W}^T \qquad [5\text{-}24]$$

2. In this situation, the equation [5-14] becomes

$$[\sqrt{1-\lambda_1}\ 0 \pm\sqrt{\lambda_3-1}]\mathbf{U} = 0 \qquad [5\text{-}25]$$

And $\mathbf{N}$ has two independent solutions, each up to a scalar, as following

$$\mathbf{N}^T = [\sqrt{1-\lambda_1}\ 0 \pm\sqrt{\lambda_3-1}]\mathbf{W}^T \qquad [5\text{-}26]$$

We know one of the solution is true but the other solution is fake. We shall call the fake one as *dual* plane.

3. In this case any $\mathbf{U}$ and hence any $\mathbf{X}$ will satisfy [5-9]. This means any $\mathbf{N}$ can be a solution. Hence $\mathbf{N}$ is undecided. Q.E.D..

In the following we sometimes use $\mathbf{N}^T$ to indicate a plane equation of the form [5-1].

Theorem 5.3 .

Let $\mathbf{N}$ be the plane of the form [5-1]. Then, the matrix $\mathbf{K}$ is rank reduced iff $\mathbf{T}^T\mathbf{R}\mathbf{N} = -1$ or iff the projection of plane [5-1] after motion is a line in the new image plane.

Proof: In step 1 and 2 we prove the first half part of the theorem and in step 3 we prove the last half of the theorem.

1. Necessary part: if **K** is rank reduced, therefore, there exists a nonzero vector $\mathbf{N}'^T$ such that

$$\mathbf{N}'^T \mathbf{K} = 0 \qquad [5\text{-}27]$$

then, from [5-2] we have

$$\mathbf{N}'^T\mathbf{X}' = \mathbf{N}'^T\mathbf{K}\ \mathbf{X} = 0 \qquad [5\text{-}28]$$

The above equation gives a line equation in the new image plane:

$$\mathbf{N}'^T [\, x'\ y'\ 1\,]^T = 0 \qquad [5\text{-}29]$$

2. Sufficient part: We prove this part by contradiction. Assume **K** has full rank, then

$$\mathbf{X} = \mathbf{K}^{-1}\mathbf{X}' \qquad [5\text{-}30]$$

Thus from [5-1] we have the new plane after motion as

$$\mathbf{N}^T\mathbf{K}^{-1}\mathbf{X}' = 1 \qquad [5\text{-}31]$$

But because the projection of the new plane is a line in the condition of the theorem, thus we have some $\mathbf{N}'$ such that

$$\mathbf{N}'^T [\, x'\ y' 1\,] = 0 \qquad [5\text{-}32]$$

then [5-28] holds. This is a contradiction to equation [5-31] since [5-32] and [5-28] will give a line not a plane equation in the space. So **K** must be rank reduced.

3. Also note that from [5-2] and [2-25] we have

$$\mathbf{R}^T\mathbf{X}' = \mathbf{X} + \mathbf{R}^T\mathbf{T} \qquad [5\text{-}33]$$

Using [5-1] we get

$$\mathbf{N}^T\mathbf{R}^T\mathbf{X}' = 1 + \mathbf{N}^T\mathbf{R}^T\mathbf{T} \qquad [5\text{-}34]$$

[5-34] has the form of [5-28] iff

$$\mathbf{N}^T\mathbf{R}^T\mathbf{T} = -1 \ \text{ or } \ \mathbf{T}^T\mathbf{R}\mathbf{N} = -1 \qquad [5\text{-}35]$$

Because in step 1 and 2 we have proved [5-28] holds iff **K** is rank reduced thus we have finished the proof of the theorem. Q.E.D..

Using the intermediate results in the proof of Theorem 5.3 we get the following Corollary:

Corollary 5.5 .

The corresponding plane equation of [5-1] after motion will be

$$\begin{cases} \mathbf{N}^T\mathbf{K}^{-1}\mathbf{X}' = \dfrac{1}{1+\mathbf{T}^T\mathbf{R}\mathbf{N}}\mathbf{N}^T\mathbf{R}^T\mathbf{X}' = 1 & \text{if } \mathbf{K} \text{ has full rank} \\ \mathbf{N}^T\mathbf{R}^T\mathbf{X}' = 0 & \text{if } \mathbf{K} \text{ is rank reduced} \end{cases} \quad \| \qquad [5\text{-}36]$$

Since in solving **K** we assume no three of four points be colinear in the space and hence in both image planes, thus if we can get a unique **K**, then that **K** must be of full rank. So we shall assume **K** be of full rank in our following discussion. So according to our discussion in section 4 and Theorem 5.1 and 5.3 we have the following theorem:

Theorem 5.4 .

Four point correspondences with no three points colinear in the space suffice to decide **K** uniquely for the plane model if both projections are not degenerate, and **K** will have full rank if it can be uniquely decided from four correspondences. ||

Before we reach a general conclusion we give theorems for each special case first.

Theorem 5.5 . (see also Tsai, [18] and Weng, [25] for different methods)

For a given **K** of the form [5-3] and if **N** is decided up to a scalar then there is a unique **R** and a unique **T** up to a scalar such that $\mathbf{K} = \mathbf{R} + \mathbf{T}\mathbf{N}^T$.

Proof: Let $\mathbf{X}_1$ and $\mathbf{X}_2$ be any two vectors such that

$$\mathbf{X}_1 \perp \mathbf{X}_2 \text{ and } \mathbf{N}^T\mathbf{X}_1 = \mathbf{N}^T\mathbf{X}_2 = 0 \qquad [5\text{-}37]$$

Then we have

$$\mathbf{R}\,\mathbf{X}_i = \mathbf{K}\,\mathbf{X}_i \triangleq \mathbf{Y}_i,\ i = 1,2 \qquad [5\text{-}38]$$

From Theorem 2.1 we know **R** can be uniquely decided by

$$\mathbf{R} = [\,\mathbf{Y}_1\,\mathbf{Y}_2\,\mathbf{Y}_1{\times}\mathbf{Y}_2\,]\,[\,\mathbf{X}_1\,\mathbf{X}_2\,\mathbf{X}_1{\times}\mathbf{X}_2\,]^{-1} \qquad [5\text{-}39]$$

After **R** is solved **T** can be solved up to a scalar by

$$\mathbf{T} = (\mathbf{K}-\mathbf{R})\mathbf{N}\,/\,\|\mathbf{N}\|^2 \qquad [5\text{-}40]$$

In this procedure we decide **R** first, then decide **T**. We can also decide **T** first by solving

$$\mathbf{W}\mathbf{K}^T\mathbf{T} = [q_1\ q_2\ q_3\ q_4\ q_5\ q_6]^T \qquad [5\text{-}41]$$

Where

$$\mathbf{W} = \begin{bmatrix} -2n_1^3+2n_1 & -2n_1^2n_2 & -2n_1^2n_3 \\ -2n_1^2n_2+n_2 & -2n_2^2n_1+n_1 & -2n_1n_2n_3 \\ -2n_1^2n_3+n_3 & -2n_1n_2n_3 & -2n_1n_3^2+n_1 \\ -2n_2^2n_1 & -2n_2^3+2n_2 & -2n_2^2n_3 \\ -2n_1n_2n_3 & -2n_2^2n_3+n_3 & -2n_2n_3^2+n_2 \\ -2n_3^2n_1 & -2n_3^2n_2 & -2n_3^3+2n_3 \end{bmatrix}, \text{ and Rank}(\mathbf{W}) = 3 \text{ always} \qquad [5\text{-}42]$$

and q_i, $i=1,2,\cdots,6$ are the elements of $\mathbf{Q}$, i.e.,

$$\mathbf{Q} = \begin{bmatrix} q_1 & q_2 & q_3 \\ q_2 & q_4 & q_5 \\ q_3 & q_5 & q_6 \end{bmatrix} = (1 - ||\mathbf{KN}||^2)\mathbf{N}\mathbf{N}^T + \mathbf{K}^T\mathbf{K} - \mathbf{I} \qquad [5\text{-}43]$$

And then $\mathbf{R}$ can be decided by

$$\mathbf{R} = \mathbf{K} - \mathbf{T}\mathbf{N}^T \qquad [5\text{-}44]$$

In [5-42] and [5-43] we assume

$$||\mathbf{N}||^2 = n_1^2+n_2^2+n_3^2 = 1 \qquad [5\text{-}45]$$

So one should first normalize $\mathbf{N}$ then solve [5-41] and then divide T by $||\mathbf{N}||$. The proof of this part is included in the Appendix A. Q.E.D..

Though the absolute value of $||\mathbf{N}||$ and $||\mathbf{T}||$ for each solution is not decided but the product $||\mathbf{N}||\,||\mathbf{T}||$ and the direction $\hat{\mathbf{N}} \triangleq \mathbf{N}/||\mathbf{N}||$ and $\hat{\mathbf{T}} \triangleq \mathbf{T}/||\mathbf{T}||$ are definite. So in the following discussion when we discuss the uniqueness of the solution we always imply that $||\mathbf{N}||\,||\mathbf{T}||$ and the direction vectors $\hat{\mathbf{N}}$ and $\hat{\mathbf{T}}$ are unique.

Theorem 5.6 . (see also Tsai ([18]) and Weng ([25]))

If all the eigenvalues of $\mathbf{K}^T\mathbf{K}$ are equal to 1, then $\mathbf{R} = \mathbf{K}$, $\mathbf{T} = 0$, and $\mathbf{N}$ is undecided.

Proof: Because all the eigenvalues of $\mathbf{K}^T\mathbf{K}$ are 1, thus there exists an orthonormal matrix $\mathbf{W}$ such that

$$\mathbf{W}^T\mathbf{K}^T\mathbf{K}\,\mathbf{W} = \mathbf{I}, \text{ or } \mathbf{K}^T\mathbf{K} = (\mathbf{W}^T)^{-1}(\mathbf{W})^{-1} = \mathbf{I} \qquad [5\text{-}46]$$

In the last equation we have used the fact that $\mathbf{W}$ is an orthonormal matrix. [5-46] tells us that $\mathbf{K}$ is an orthonormal matrix. Now consider a triad constructed by three non-zero unit vectors $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$ such that

$$\mathbf{X}_1 \perp \mathbf{X}_2,\ \ \mathbf{N} \perp \mathbf{X}_i\ ,\ i = 1{,}2, \text{ and } \mathbf{X}_3 = \mathbf{X}_1 \times \mathbf{X}_2\,. \qquad [5\text{-}47]$$

Because of [5-46] and [5-47] we have

$$(\mathbf{KX}_i)^T(\mathbf{KX}_j) = \delta_{ij},\ \ i,j = 1, 2, 3 \qquad [5\text{-}48]$$

where δ_{ij} is a Dirac function. [5-48] means $\mathbf{KX}_i$, i = 1, 2, 3 construct a triad either. Thus we must have

$$\mathbf{KX}_3 = \alpha(\mathbf{KX}_1 \times \mathbf{KX}_2) \qquad [5\text{-}49]$$

for some constant α. Still because of the orthonormality of $\mathbf{K}$ and [5-47] we should have

$$\|\mathbf{KX}_3\| = \|\mathbf{X}_3\| = \|\mathbf{X}_1\|\,\|\mathbf{X}_2\| \qquad [5\text{-}50]$$

and from [5-49] and [5-48] we get

$$|\alpha|\,\|\mathbf{KX}_1 \times \mathbf{KX}_2\| = |\alpha|\,\|\mathbf{KX}_1\|\,\|\mathbf{KX}_2\| = |\alpha|\,\|\mathbf{X}_1\|\,\|\mathbf{X}_2\| = \|\mathbf{X}_1\|\,\|\mathbf{X}_2\| \qquad [5\text{-}51]$$

Thus we have $\alpha = \pm 1$. Since a rigid motion should not change the rigdity of the triad, by requiring keeping the chirality we must have $\alpha = 1$. Hence we have

$$\mathbf{KX}_1 \times \mathbf{KX}_2 = \mathbf{K}(\mathbf{X}_1 \times \mathbf{X}_2) \qquad [5\text{-}52]$$

Because $\mathbf{X}_i$, i = 1,2, are perpendicular to $\mathbf{N}^T$, thus we have

$$\mathbf{RX}_i = \mathbf{KX}_i,\ \ i=1{,}2 \qquad [5\text{-}53]$$

and then from [2-28], [5-53] and [5-52] we get

$$\mathbf{RX}_3 = \mathbf{R}(\mathbf{X}_1 \times \mathbf{X}_2) = \mathbf{RX}_1 \times \mathbf{RX}_2 = \mathbf{KX}_1 \times \mathbf{KX}_2 = \mathbf{K}(\mathbf{X}_1 \times \mathbf{X}_2) \qquad [5\text{-}54]$$

Finally from [5-54] and [5-53] we reach

$$\mathbf{R}\,[\,\mathbf{X}_1\,\mathbf{X}_2\,\mathbf{X}_1 \times \mathbf{X}_2\,] = [\,\mathbf{KX}_1\,\mathbf{KX}_2\,\mathbf{KX}_1 \times \mathbf{KX}_2\,] = \mathbf{K}\,[\,\mathbf{X}_1\,\mathbf{X}_2\,\mathbf{X}_1 \times \mathbf{X}_2\,] \qquad [5\text{-}55]$$

Since $[\,\mathbf{X}_1\,\mathbf{X}_2\,\mathbf{X}_1 \times \mathbf{X}_2\,]$ has three orthogonal columns and is hence invertible, we have got $\mathbf{R} = \mathbf{K}$. And then by $\mathbf{TN}^T = \mathbf{K} - \mathbf{R} = 0$ and $\mathbf{N}^T \neq 0$ we have $\mathbf{T} = 0$. But $\mathbf{N}$ is undecided. Q.E.D..

Note that the cases in Theorem 5.2 are incompatible, from Theorem 5.2 to Theorem 5.6 we immediately have the following general theorem.

Theorem 5.7 . (see also Tsai, [18] and Weng, [25] for different statements)

For any matrix $\mathbf{K}$ of the form [5-3],

1. iff all the three eigenvalues of $\mathbf{K}^T\mathbf{K}$ are 1 then $\mathbf{R} = \mathbf{K}$, and $\mathbf{T} = 0$.

2. iff two of the three eigenvalues of $\mathbf{K}^T\mathbf{K}$ are 1 then $\mathbf{R}$ can be uniquely decided and $\mathbf{N}$ and $\mathbf{T}$ can be uniquely decided to within a scalar.

3. iff all the three eigenvalues of $\mathbf{K}^T\mathbf{K}$ are different then there are two independent sets of solutions for $\mathbf{R}$, $\mathbf{N}$, and $\mathbf{T}$, where $\mathbf{N}$ and $\mathbf{T}$ can only be decided to within a scalar. ∎

So far we have known that if there is no translation, then all the three eigenvalues of $\mathbf{K}^T\mathbf{K}$ will be 1 and the motion can be uniquely decided but the plane cannot be decided. But in what cases $\mathbf{K}^T\mathbf{K}$ will have two identical eigenvalues and in what cases it has three different eigenvalues? That is, In what situation the motion and the plane can be uniquely decided and in what situation they cannot? How are the eigenvalues of $\mathbf{K}^T\mathbf{K}$ associated with the motion and the plane position? In case two solutions exist how are the dual plane and motion related to the real plane and motion? The following theorem answers these questions. To make the statement brief we use the convention that a zero vector is said to be parallel to any vector. And we generally assume there exist two sets of solutions for [5-3]. If the the dual solution is identical to the real solution then we have the unique solution.

Theorem 5.8 .

For any matrix $\mathbf{K}$ of the form [5-3], let $\mathbf{R}$, $\mathbf{T}$, and $\mathbf{N}$ be the real solution, then the dual plane $\mathbf{N}_d$, dual translation vector $\mathbf{T}_d$ and dual rotation matrix $\mathbf{R}_d$ satisfy

$$\mathbf{N}_d = \alpha(\mathbf{N}\,\|\mathbf{T}\|^2 + 2\mathbf{R}^T\mathbf{T}) = \alpha(\mathbf{K}^T + \mathbf{R}^T)\mathbf{T} \qquad [5\text{-}56]$$

$$\mathbf{T}_d = \beta(\mathbf{T}\,\|\mathbf{N}\|^2 + 2\mathbf{R}\mathbf{N}) = \beta(\mathbf{K} + \mathbf{R})\mathbf{N} \qquad [5\text{-}57]$$

$$\mathbf{R}_d = \mathbf{K} - \alpha\beta(\mathbf{K} + \mathbf{R})\mathbf{N}\mathbf{T}^T(\mathbf{K} + \mathbf{R}) \qquad [5\text{-}58]$$

where α, β are constants and $\alpha\beta$ is to be decided by making

$$\mathbf{R}_d^T\mathbf{R}_d = [\mathbf{K} - \alpha\beta(\mathbf{K} + \mathbf{R})\mathbf{N}\mathbf{T}^T(\mathbf{K} + \mathbf{R})]^T[\mathbf{K} - \alpha\beta(\mathbf{K} + \mathbf{R})\mathbf{N}\mathbf{T}^T(\mathbf{K} + \mathbf{R})] = \mathbf{I} \qquad [5\text{-}59]$$

And there are still many other relations can be derived from [5-56] and [5-57], e.g.

$$(\mathbf{R}_d^T\mathbf{T}_d + \mathbf{R}^T\mathbf{T}\,\|\mathbf{T}_d\|^2) \;/\!/\; \mathbf{N} \qquad [5\text{-}60]$$

So iff $\mathbf{R}^T\mathbf{T} \,/\!/\, \mathbf{N}$ then $\mathbf{K}^T\mathbf{K}$ has three identical eigenvalues and the dual solution is identical to the real solution aside from a scalar in $\mathbf{N}$ and $\mathbf{T}$, in other words, $\mathbf{R}$, $\mathbf{T}$ and $\mathbf{N}$ has unique solution.

Proof: What we need to show is that iff $\mathbf{R}^T\mathbf{T} \,/\!/\, \mathbf{N}$ then the dual plane is identical to the real plane aside from a scalar. Assume any vector $\mathbf{X}$ such that

$$\mathbf{N}^T\mathbf{X} \neq 0 \qquad [5\text{-}61]$$

but

$$\mathbf{X}^T\,\mathbf{K}^T\mathbf{K}\,\mathbf{X} = \mathbf{X}^T\mathbf{X}, \qquad [5\text{-}62]$$

Expending the left side of [5-61] and rearranging the resulting formula leads to

$$\mathbf{X}^T(\mathbf{R}^T\mathbf{T}\mathbf{N}^T + \mathbf{N}\mathbf{T}^T\mathbf{R} + \mathbf{N}\mathbf{T}^T\mathbf{T}\mathbf{N}^T)\,\mathbf{X} = 0, \qquad [5\text{-}63]$$

$$\text{or} \quad \mathbf{X}^T(\mathbf{N}\,\|\mathbf{T}\|^2 + 2\mathbf{R}^T\mathbf{T})\cdot\mathbf{N}^T\mathbf{X} = 0$$

$$\text{or} \quad \mathbf{X}^T(\mathbf{N}\,\|\mathbf{T}\|^2 + 2\mathbf{R}^T\mathbf{T}) = 0 \qquad [5\text{-}64]$$

because of [5-61]. Obviously (see Theorem 5.2), [5-64] defines the dual plane, in other words, $\mathbf{N}_d \,/\!/\, (\mathbf{N}\,\|\mathbf{T}\|^2 + 2\mathbf{R}^T\mathbf{T})$, thus we get [5-56]. By the *duality property* we immediately have [5-57]. And [5-58] results because $\mathbf{K} = \mathbf{R}_d + \mathbf{T}_d\mathbf{N}^T{}_d$. [5-59] is a requirement of the orthonormality of $\mathbf{R}_d$. Since a scalar in $\mathbf{N}_d$ will not affect $\hat{\mathbf{N}}_d$, $\hat{\mathbf{T}}_d$, and $\|\mathbf{T}_d\|\,\|\mathbf{N}_d\|$ so without losing generality we select one solution is

$$\mathbf{N}_d = \mathbf{N}\,\|\mathbf{T}\|^2 + 2\mathbf{R}^T\mathbf{T} \qquad [5\text{-}65]$$

Similarly we can consider $\mathbf{N}$ is a dual plane of $\mathbf{N}_d$, thus, just like [5-56] we should have

$$\mathbf{N} = \gamma(\mathbf{N}_d\,\|\mathbf{T}_d\|^2 + 2\mathbf{R}_d^T\mathbf{T}_d) \qquad [5\text{-}66]$$

where γ is a non-zero constant because $\mathbf{N}$ cannot be zero. Replace [5-65] into [5-66] we get

$$(1-\gamma\,\|\mathbf{T}\|^2\,\|\mathbf{T}_d\|^2)\mathbf{N} = 2\gamma(\mathbf{R}^T\mathbf{T}\,\|\mathbf{T}_d\|^2 + 2\mathbf{R}_d^T\mathbf{T}_d) \qquad [5\text{-}67]$$

So it is clear that $(\mathbf{R}_d^T\mathbf{T}_d + \mathbf{R}^T\mathbf{T}\,\|\mathbf{T}_d\|^2)$ is parallel to $\mathbf{N}$. From [5-56] we know iff $\mathbf{R}^T\hat{\mathbf{T}} \,/\!/\, \mathbf{N}$ then $\mathbf{N}_d$ is parallel to $\mathbf{N}$, and $\mathbf{N}_d$ is identical to $\mathbf{N}$ aside from a scalar. Thus we have proved that iff $\mathbf{R}^T\hat{\mathbf{T}} \,/\!/\, \mathbf{N}$ then $\mathbf{K}$ has unique decomposition into $\mathbf{R} + \mathbf{T}\mathbf{N}^T$. Q.E.D..

Now the problem that remains is how we figure out the spurious solution in case two solutions exist. There are many methods, such as multiple plane patches, multiple views and *motion epipolar* constraint [2-39] etc., which can do the job. We shall only discuss the last two methods here. Tsai etc. ([19]) investigated three view problem earlier and established the uniqueness theorem in a long paper. Unfortunately it seems to us that the proof and the conclusion is only true for a special case, as verified by our following theorem and many numerical results. Their theorem and proof are based on one assumption which says that iff the plane and the dual plane solved from t_2 to t_1 are excatly the same as those solved from t_2 to t_3 then the motion from t_2 to t_1 and from t_2 to t_3 are both undecided, but in this case, they proved that the motion from $t_1$tot_3 is a pure rotation and is uniquely decided. We would like to point out that if the plane and the dual plane solved from t_2 to t_1 are not excatly the same as but are just parallel to those solved from t_2 to t_3 separately, then the motion is also undecided, because in monocular vision we can only decide the plane up to a scalar and hence cannot distinguish two planes of the same orientation. And in this case, the motion from t_1 to t_3 is not a pure rotation in general and hence cannot be uniquely decided either. Our conclusion has been verified by many numerical results.

Lemma 5.1 .

If $\mathbf{K}^{-1}$ exists, then vector $\mathbf{N}^T$ is parallel to $\mathbf{N}_d{}^T$ iff $\mathbf{N}^T\mathbf{K}^{-1}$ is parallel to $\mathbf{N}_d{}^T\mathbf{K}^{-1}$, and $\mathbf{N}^T = 0$ iff $\mathbf{N}^T\mathbf{K}^{-1} = 0$.

Proof: We only prove the sufficient part, the necesary part can be proved the same way. If $\mathbf{N}$ is parallel to $\mathbf{N}_d$, then there exists some number α such that $\mathbf{N} = \alpha\mathbf{N}_d$. Then we have $\mathbf{N}\mathbf{K}^{-1} = \alpha\mathbf{N}_d\mathbf{K}^{-1}$. This means $\mathbf{N}\mathbf{K}^{-1}$ is parallel to $\mathbf{N}_d\mathbf{K}^{-1}$. And if $\mathbf{N}^T = 0$, it is obvious $\mathbf{N}^T\mathbf{K}^{-1} = 0$ Q.E.D..

Lemma 5.2 .

If the plane $\mathbf{N} \neq 0$, then the dual plane $\mathbf{N}_d = 0$ iff the translation $\mathbf{T} = 0$.

Proof: From Theorem 5.8 we know the dual plane is represented by

$$\mathbf{N}_d = \alpha(\mathbf{N} \,||\mathbf{T}||^2 + 2\mathbf{R}^T\mathbf{T}) = (\mathbf{K}^T + \mathbf{R}^T)\mathbf{T} \qquad [5\text{-}68]$$

hence if $\mathbf{T} = 0$ then $\mathbf{N}_d$=0. if $\mathbf{N}_d = 0$ then we see $\mathbf{R}_d = \mathbf{K}$. This means $\mathbf{k}$ is a rotation matrix. From Theorem 5.6 we know $\mathbf{T} = 0$. Q.E.D..

In fact when $\mathbf{T} = 0$, $\mathbf{N}$ and $\mathbf{N}_d$ are all undecided. This lemma just shows that unless $\mathbf{T} = 0$ then a zero vector can never be the solution of the plane.

Theorem 5.9 .

Let

$$\mathbf{X}' = (\mathbf{R}_1 + \mathbf{T}_1\mathbf{N}_1^T)\mathbf{X} = \mathbf{K}_1\mathbf{X} \quad \text{with } \mathbf{N}_1^T\mathbf{X} = 1 \qquad [5\text{-}69]$$

$$\mathbf{X}'' = (\mathbf{R}_2 + \mathbf{T}_2\mathbf{N}_2^T)\mathbf{X}' = \mathbf{K}_2\mathbf{X}' \quad \text{with } \mathbf{N}_2^T\mathbf{X}' = 1 \qquad [5\text{-}70]$$

and

$$\mathbf{X}'' = (\mathbf{R}_3 + \mathbf{T}_3\mathbf{N}_1^T)\mathbf{X} = \mathbf{K}_3\mathbf{X} \qquad [5\text{-}71]$$

with

$$\mathbf{R}_3 = \mathbf{R}_2\mathbf{R}_1, \quad \mathbf{T}_3 = \mathbf{T}_2 + \mathbf{R}_2\mathbf{T}_1. \qquad [5\text{-}72]$$

where $\mathbf{X}, \mathbf{X}', \mathbf{X}''$ are the same scene point at time t_1, t_2 and t_3 separately. Assume all $\mathbf{K}_i$,i = 1,2,3, can be solved uniquely by four correspondences and no two views alone can decide the motion uniquely, then iff there exists no number α such that

$$(\mathbf{K}_2^T + \mathbf{R}_2^T)\mathbf{T}_2 = \alpha(\mathbf{K}_1^T)^{-1}(\mathbf{K}_1^T + \mathbf{R}_1^T)\mathbf{T}_1 \qquad [5\text{-}73]$$

then three views suffice to decide the motions uniquely, and the planes can be d ecided uniquely either if any one of the translation vectors is not zero.

Proof: There are three motions in three view: from t_1 to t_2, from t_2 to t_3, and from t_1 to t_3. We know that in each motion, there are at most two sets of solutions. Thus, if one motion is uniquely decided, then all other motions can be consequently decided either. And so do the planes.

Now let's prove the theorem. Because we assume no two views alone can decide the motion uniquely, thus each $\mathbf{K}_i^T\mathbf{K}_i$, i = 1,2,3, has full rank, three different eigenvalues and hence two sets of decompositions. What we need to find is the condition under which the dual solutions happen to make up a consistent motion. According to Theorem 5.8, from $\mathbf{K}_1$ we can get one of the dual plane of $\mathbf{N}_1$ as

$$\mathbf{N}_{1d} = (\mathbf{K}_1^T + \mathbf{R}_1^T)\mathbf{T}_1, \qquad [5\text{-}74]$$

from $\mathbf{K}_2$ we can get one of the dual plane of $\mathbf{N}_2$ as

$$\mathbf{N}_{2d} = (\mathbf{K}_2^T + \mathbf{R}_2^T)\mathbf{T}_2, \qquad [5\text{-}75]$$

and from $\mathbf{K}_3$ we can get another of the dual plane of $\mathbf{N}_1$ as

$$\mathbf{N}'_{1d} = (\mathbf{K}_3^T + \mathbf{R}_3^T)\mathbf{T}_3, \qquad [5\text{-}76]$$

From Corollary 5.5 we have

$$\mathbf{N}_2 = \frac{1}{1+\mathbf{T}_1^T\mathbf{R}_1\mathbf{N}_1}\mathbf{RN} = (\mathbf{K}^T)^{-1}\mathbf{N}_1 \qquad [5\text{-}77]$$

From Theorem 5.7 and Theorem 5.8 we know that $\mathbf{N}_{1d}$ and $\mathbf{N}'_{1d}$ cannot make up a consistent motion with $\mathbf{N}_2$ otherwise $\mathbf{N}_{1d}$ and $\mathbf{N}'_{1d}$ must be parallel to $\mathbf{N}_1$, which means that $\mathbf{K}_1$ and $\mathbf{K}_3$ have unique decompositions and hence contradicts our assumption. Similarly from [5-77] we know $\mathbf{N}_{2d}$ cannot make up a consistent motion with $\mathbf{N}_1$ since N_2 is uniquely defined by [5-77], that is, given $\mathbf{K}_1$, the plane after motion is definitely decided. Thus what we need to find is just the condition that $\mathbf{N}_{1d}$, $\mathbf{N}_{2d}$, $\mathbf{N}'_{1d}$ make up a consistent motion. Again from Corollary 5.5 we know, to make this happen, we need and only need $\mathbf{N}'_{1d} = \beta\mathbf{N}_{1d}$, $\mathbf{K}_1^T\mathbf{N}_{2d} = \alpha\mathbf{N}_{1d}$ for some constants alpha, and β. Thus we get the following equations:

$$\mathbf{K}_1^T(\mathbf{K}_2^T + \mathbf{R}_2^T)\mathbf{T}_2 = \alpha(\mathbf{K}_1^T + \mathbf{R}_1^T)\mathbf{T}_1 \qquad [5\text{-}78]$$

$$(\mathbf{K}_3^T + \mathbf{R}_3^T)\mathbf{T}_3 = \beta(\mathbf{K}_1^T + \mathbf{R}_1^T)\mathbf{T}_1 \qquad [5\text{-}79]$$

Replacing [5-72] into [5-79] we get

$$(\mathbf{K}_3^T + \mathbf{R}_3^T)(\mathbf{T}_2 + \mathbf{R}_2\mathbf{T}_1) = \beta(\mathbf{K}_1^T + \mathbf{R}_1^T)\mathbf{T}_1 \qquad [5\text{-}80]$$

or

$$(\mathbf{K}_1^T\mathbf{K}_2^T\mathbf{T}_2^T + \mathbf{K}_1^T\mathbf{K}_2^T\mathbf{R}_2\mathbf{T}_1 + \mathbf{R}_3^T\mathbf{T}_2 + \mathbf{R}_3^T\mathbf{R}_2^T\mathbf{T}_1) = \beta(\mathbf{K}_1^T + \mathbf{R}_1^T)\mathbf{T}_1 \qquad [5\text{-}81]$$

It turns out iff

$$\beta = 1 + \alpha \qquad [5\text{-}82]$$

then [5-79] and [5-78] can both hold. Therefore given an α there is exactly one β such that [5-78] and [5-79] both hold. So what we need to know is the condition under which [5-78] holds. Rewriting [5-78] we have

$$\mathbf{R}_1\mathbf{N}_1\|\mathbf{T}_2\|^2/(1+\mathbf{T}_1^T\mathbf{R}_1\mathbf{N}_1) + 2\mathbf{R}_2^T\mathbf{T}_2 = \alpha(\mathbf{K}_1^T)^{-1}(\mathbf{K}_1^T + \mathbf{R}_1^T)\mathbf{T}_1 \qquad [5\text{-}83]$$

So from [5-82] we see, in general cases for any fixed $\mathbf{R}_1$, $\mathbf{T}_1$, $\mathbf{N}_1$ and α, there are an infinite number of solutions for $\mathbf{R}_2$, $\mathbf{T}_2$ such that [5-78] is satisfied. So in general cases we cannot assure that three views can decide the motion and the plane uniquely. However, for a given $\mathbf{R}_2$,$\mathbf{T}_2$, there might be no α such that [5-78] is satisfied. if we consider that $\mathbf{R}_2$ and $\mathbf{T}_2$ are totally random, then the probability with which [5-78] happens to be met is almost zero. So we have proved the theorem. Q.E.D..

However, we find if $\alpha = -1$ or $\beta = 1$ in Theorem 5.9, then we find we will get the same conclusion as Tsai did. We state this fact as a corollary.

Corollary 5.6

We use the same notations as in Theorem 5.9. If it happens that in [5-73] $\alpha = -1$, the motion from t_1 to t_3 can be uniquely decided; or in [5-79] if $\beta = 1$ then the motion from $t_1$tot_2 can be decided uniquely.

Proof: If $\alpha = -1$, from [5-82] we know $\beta = 0$, then from [5-79] and [5-76] we know $\mathbf{N}'_{1d} = 0$. Hence according to Lemma 5.2 we know $\mathbf{T}_3 = 0$ and $\mathbf{R}_3 = \mathbf{K}_3$. Thus the motion from t_1 to t_3 can be uniquely decided. Simmilarly one can prove the other part conclusion when $\beta = 1$. Q.E.D..

From Theorem 5.9 we know that we cannot rely on three views to figure out the fake solution, though multiple view matching of course gives more evidences of the true motion and plane. However we still have the *motion epipolar* line constraint to judge the correct one. Suppose in the scene there are many other points not on the same plane we are solving, and if the correspondeces of these points are available then we can check on which *motion epipolar* line defined by the two sets of motions as in form of [2-39] their correspondences lie. And then we can decide the right motion. However, if all the scene points fail to satisfy the "surface condition" defined in [22], then this technique may sometimes again break down. Fortunately the probability with which this situation happens in the nature is also zero.

Now we come to the problem how we find out the four correspondences in a plane. In the stereo case or when range data available, the plane equation can be solved before one solve the motion. This problem is almost equivalent to the three point correspondence technique described by Theorem 2.2. For the monocular case where depth information is not available before solving the motion, then it seems this technique can hardly be used in application. However, by playing a trick and searching in a reasonable window we will find we only need three point correspondences to solve the motion, but we still need many other correspondences to judge the correct one. The algorithm will be discussed in next section. But we first give three useful theorems here.

Theorem 5.10 .

If the motion **R** and **T** is solved and $\mathbf{T} \neq 0$, then given any three point coorespondences not colinear in space, one can find the plane they lie uniquely.

Proof: [5-2] can be rewritten as

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} \frac{Z'}{Z} = [\, \mathbf{R} + \mathbf{T}\mathbf{N}^T \,] \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \qquad [5\text{-}84]$$

Thus one correspondence gives two dependent equations for $\mathbf{N} = [\, n_1 \; n_2 \; \text{nsub3} \,]$:

$$x' = \frac{(r_{11}+t_1n_1)x + (r_{12}+t_1n_2)y + (r_{13}+t_1n_3)}{(r_{31}+t_3n_1)x + (r_{32}+t_3n_2)y + (r_{33}+t_3n_3)} \qquad [5\text{-}85]$$

$$y' = \frac{(r_{21}+t_2n_1)x + (r_{22}+t_2n_2)y + (r_{23}+t_2n_3)}{(r_{31}+t_3n_1)x + (r_{32}+t_3n_2)y + (r_{33}+t_3n_3)} \qquad [5\text{-}86]$$

Reordering the above equations we get

$$xn_1 + yn_2 + n_3 = \frac{(r_{11}-r_{31}x')x + (r_{12}-r_{32}x')y + r_{13}-r_{33}x'}{t_3x' - t_1} \qquad [5\text{-}87]$$

or

$$xn_1 + yn_2 + n_3 = \frac{(r_{21}-r_{31}y')x + (r_{22}-r_{32}y')y + r_{23}-r_{33}y'}{t_3y' - t_2} \qquad [5\text{-}88]$$

[5-87] and [5-88] should be the same. It is easy to know that three non-colinear points in space will suffice to solve **N** from [5-87] or [5-88]. Q.E.D..

[5-87] relies on x' and [5-88] relies on y'. One can selectively use [5-87] or [5-88] if he feels that x' or y' may be more accurate. The next Theorem tells us that if we know the plane that three non-colinear points lie but not the plane that their correspondences lie then we may get up to four sets of solutions. Again, by multiple view matching, motion epipolar line constraint, we may get the unique solution.

Theorem 5.11 .

Three non-colinear point correspondences with depths or the plane equation in one image known, will give up to four sets of motion solution.

Proof: In this theorem we shall reverse the convention on the use of prime to make notation simpler. Let's consider three noncolinear points $\mathbf{X}'_i$, i = 1,2,3, and their correspondences $\mathbf{X}_i$, i =1,2,3, and assume they are related by

$$\mathbf{X}_i = \mathbf{R}\mathbf{X}'_i + \mathbf{T}, \quad i = 1,2,3 \qquad [5\text{-}89]$$

Because the plane equation or the depths in $\mathbf{X}'_i$, i =1,2,3, are known, we can assume $\mathbf{X}'_i$, i =1,2,3, are all known. Thus we have

$$\mathbf{X}_1 - \mathbf{X}_2 = \mathbf{R}(\mathbf{X}'_1 - \mathbf{X}'_2), \quad \mathbf{X}_1 - \mathbf{X}_3 = \mathbf{R}(\mathbf{X}'_1 - \mathbf{X}'_3) \qquad [5\text{-}90]$$

Becasue of the noncolinearity of $\mathbf{X}'_i$, i = 1,2,3, [5-90] gives two independent matrix equations. Now using [2-19], we shall have following equations from [5-89]:

$$(Z_1\mathbf{\Theta}_1 - Z_i\mathbf{\Theta}_i)\cdot(Z_1\mathbf{\Theta}_1 - Z_j\mathbf{\Theta}_j) = (\mathbf{X}'_1 - \mathbf{X}'_i)\cdot(\mathbf{X}'_1 - \mathbf{X}'_j), \quad i,j = 2,3 \qquad [5\text{-}91]$$

let

$$\lambda_1 = \frac{Z_2}{Z_1}, \quad \lambda_2 = \frac{Z_3}{Z_1} \qquad [5\text{-}92]$$

$$\rho_1 = \|\mathbf{X}'_1 - \mathbf{X}'_2\|^2, \quad \rho_2 = \|\mathbf{X}'_1 - \mathbf{X}'_3\|^2, \quad \rho_3 = (\mathbf{X}'_1 - \mathbf{X}'_2)\cdot(\mathbf{X}'_1 - \mathbf{X}'_3)$$

then [5-91] gives following three equations:

$$\lambda_1^2(x_2^2+y_2^2+1) - 2\lambda_1(x_1x_2+y_1y_2+1) + (x_1^2+y_1^2+1) = \rho_1 / Z_1^2 \qquad [5\text{-}93]$$

$$\lambda_2^2(x_3^2+y_3^2+1) - 2\lambda_1(x_1x_3+y_1y_3+1) + (x_1^2+y_1^2+1) = \rho_2 / Z_1^2 \qquad [5\text{-}94]$$

$$\lambda_1\lambda_2(x_2x_3+y_2y_3+1) - \lambda_1(x_2x_1+y_2y_1+1) - \lambda_2(x_1x_3+y_1y_3+1) + (x_1^2+y_1^2+1) = \rho_3 / Z_1^2 \qquad [5\text{-}95]$$

Combining [5-93] and [5-95] and deleting Z_1 gives

$$\lambda_1^2(x_2^2+y_2^2+1)\rho_3 - \lambda_1(x_1x_2+y_1y_2+1)(2\rho_3-\rho_1) + \lambda_2(x_1x_3+y_1y_3+1)\rho_1$$
$$- \lambda_1\lambda_2(x_2x_3+y_2y_3+1)\rho_1 + (\rho_3-\rho_1)(x_1^2+y_1^2+1) = 0 \qquad [5\text{-}96]$$

or

$$a_0\lambda_1^2 + a_1\lambda_1 + a_2\lambda_2 + a_3\lambda_1\lambda_2 + a_4 = 0 \qquad [5\text{-}97]$$

and combining [5-94] and [5-95] and deleting Z_1 gives

$$\lambda_2^2(x_3^2+y_3^2+1)\rho_3 - \lambda_2(x_1x_3+y_1y_3+1)(2\rho_3-\rho_2) + \lambda_1(x_1x_2+y_1y_2+1)\rho_2$$
$$- \lambda_1\lambda_2(x_2x_3+y_2y_3+1)\rho_2 + (\rho_3-\rho_2)(x_1^2+y_1^2+1) = 0 \qquad [5\text{-}98]$$

or

$$b_0\lambda_1^2 + b_1\lambda_1+b_2\lambda_2+b_3\lambda_1\lambda_2 +b_4 = 0 \qquad [5\text{-}99]$$

From [5-97] we can get

$$\lambda_2 = -\frac{a_0\lambda_1^2+a_1\lambda_1+a_4}{a_2+a_3\lambda_1} \qquad [5\text{-}100]$$

then replacing [5-100] into [5-99] will give a fourth order polynomial equation for λ_1. So we have at most four real, positive solutions for λ_1. After λ_1 is solved, λ_2 and Z_1 will be given by [5-100] and [5-93] etc.. Thus Z_1, Z_2, and Z_3 will have up to four sets of solutions. From Theorem 2.2 we know, **R** and **T** will also have up to four sets of solutions. We can require λ_1, λ_2, and Z_1 be positive to figure out some fake solutions, but we cannot assure this always makes the solution unique. Q.E.D..

To figure out the spurious solutions we may again use the motion epipolar constraint. But if the motion is small, we can also solve [2-17] to get an approximate solution of the motion and then use Fact 2.4, i.e. [2-59], to pick out the real solution.

From Corollary 5.5, Theorem 5.10 and Theorem 5.11 we see oringinally we may need 4 plane point correspondences to solve motion, but sequentially we only need three point correspondences to solve motion. Since three non-colinear points always lie in a plane, thus this method can be extended to non-planar surfaces. In fact Theorem 5.11 holds not only for planar surfaces if only depths are known in one image. Sequentially, relative depth information may be got from previous motion estimation but may not be always available. When translation is zero, then the depth information may not be given. Another problem is that if one uses the previous information to solve following-on motion, the errors may accumulate. That is, the error in the estimation of previous motion may pass to the following motion estimation. So we have to have a restart algorithm to correct the wrong estimations. A realistic algorithm using only three point correspondences is discussed in Section 8.

6. Traditional Two Step 8-point Linear Algorithm

In this section we investigate the advantages and disadvantages of the traditional 8-point linear algorithm and then propose an improved one. The traditional 8-point algorithm assumes the translation $\mathbf{T} \neq 0$. Then at least 8-point correspondences satisfying the surface assumption ([22], Zhuang) are required to give a solution of $\mathbf{T} \times \mathbf{R}$ by solving [2-39]. Let

$$\mathbf{G} = \mathbf{T}\times = \begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{g}_1^T \\ \mathbf{g}_2^T \\ \mathbf{g}_3^T \end{bmatrix} \qquad [6\text{-}1]$$

$$\mathbf{E} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix} = \mathbf{T}\times\mathbf{R} = \begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{bmatrix} \mathbf{R} = \mathbf{G}\ \mathbf{R} \qquad [6\text{-}2]$$

Then [2-39] becomes

$$[\, x'\ y'\ 1\,]\ \mathbf{E}\ [\, x\ y\ 1\,]^T = 0 \qquad [6\text{-}3]$$

and n point correspondences will give the following equation

$$\mathbf{A}\ \mathbf{h} = 0 \qquad [6\text{-}4]$$

where

$$\mathbf{A} = \begin{bmatrix} x'_1x_1 & x'_1y_1 & x'_1 & y'_1x_1 & y'_1y_1 & y'_1 & x_1 & y_1 & 1 \\ & & \cdot\cdot & \cdot & \cdot & \cdot & \cdot & & \\ & & & & \cdot & & & & \\ & & & & \cdot & & & & \\ x'_nx_n & x'_ny_n & x'_n & y'_nx_n & y'_ny_n & y'_n & x_n & y_n & 1 \end{bmatrix}$$

$$\mathbf{h} = [\, h_1\ h_2\ \cdots\ h_9\,]^T \qquad [6\text{-}5]$$

Then the traditional two step linear algorithm first gets a least square solution of $\mathbf{E}$ from [6-4] and then recovers $\mathbf{T}$ and $\mathbf{R}$ from $\mathbf{E}$. The least square solution of [6-4] is just the eigenvector of unit norm of $\mathbf{A}^T\mathbf{A}$ associated with the smallest eigenvalue of $\mathbf{A}^T\mathbf{A}$. There also exist several linear algorithms for recovering $\mathbf{T}$ and $\mathbf{R}$ from $\mathbf{E}$.

The advantage of this algorithm is the linearity and closedness of the solution as well as the integrability of many correspondences. But it also has some disadvantages:

1. it cannot handle the case where $\mathbf{T} = 0$ or $\mathbf{T}$ is small, that is, one cannot compare which estimation is better: a pure rotation or a rotation plus a small translation. And when the translation is very small, to recover $\mathbf{T}$ and $\mathbf{R}$ from $\mathbf{E}$ might be extremely erroneous.

equivalent to having a nonzero solution $\mathbf{E}$ such that $\mathbf{E}$ is not equal to $\alpha\mathbf{T}\times\mathbf{R}$ for any nonzero α. ||

We would like to point out that this is true only when there is no noise. When noise exists in the matrix $\mathbf{A}$ then even [6-4] is not degenerate, not every solution $\mathbf{E}$ from [6-4] can be decomposed into $\alpha\mathbf{T}\times\mathbf{R}$, as we will see in the next subsection.

Surface Assumption

The group of points S used in $\mathbf{A}$ does not lie on a quadratic surface of the following form:

$$[X\ Y\ Z]\mathbf{U}[X\ Y\ Z]^T + \mathbf{T}^T\mathbf{R}\mathbf{U}[X\ Y\ Z]^T = 0 \qquad [6\text{-}6]$$

with

$$||\mathbf{U} + \mathbf{U}^T|| + ||\mathbf{T}^T\mathbf{R}\mathbf{U}|| \neq 0 \qquad [6\text{-}7]$$

where $\mathbf{U}$ can be any 3×3 matrix. ||

This assumption is only a sufficient condition for the uniqueness of motion solution but not necessary. In case of noise, this condition still cannot guarantee the solution $\mathbf{E}$ from [6-4] can be decomposed in to [6-2].

Theorem

Under the Surface Assumption, the two-view motion equation [6-4] has a rank 8 and a general solution $\alpha\mathbf{T}\times\mathbf{R}$, when $\mathbf{T} \neq 0$, or a rank 6 and a general solution $\mathbf{T}_1\times\mathbf{R}$, where $\mathbf{T}_1$ is any (nonzero, added by the authors) real vector, when $\mathbf{T} = 0$. ||

Still we would like to point out that this theorem is only true in case of no noise. The reason we mention this surface assumption here is that this assumption must be satisfied to apply the 8-point linear algorithm. However, even it is satisfied we may still not find a solution $\mathbf{E}$ of the form [6-2] from [6-4] in case of noise. This is to be seen in the following.

6.2. Conditions and Algorithm for Decomposing $\mathbf{E}$ into $\mathbf{T}\times\mathbf{R}$

Based on Theorem 2.1, we propose our algorithm to recover $\mathbf{T}$ and R from $\mathbf{E}$. If a matrix $\mathbf{E}$ has a decomposition of the form [6-4] with $\mathbf{T} \neq 0$ we say that $\mathbf{E}$ has a motion decomposition. We shall first state some necessary conditions for a matrix $\mathbf{E}$ to have a motion decomposition in theorems, then propose our algorithm to do the decomposition, and at last give the necessary and sufficient condition for a matrix $\mathbf{E}$ to have a motion decomposition. Let's transpose [6-2] into the following form:

$$\mathbf{E}^T = [\,\mathbf{e}_1\,\mathbf{e}_2\,\mathbf{e}_3\,] = \begin{bmatrix} h_1 & h_4 & h_7 \\ h_2 & h_5 & h_8 \\ h_3 & h_6 & h_9 \end{bmatrix} = \mathbf{R}^T \begin{bmatrix} 0 & t_3 & -t_2 \\ -t_3 & 0 & t_1 \\ t_2 & -t_1 & 0 \end{bmatrix} = \mathbf{R}^T[\,\mathbf{g}_1\,\mathbf{g}_2\,\mathbf{g}_3\,] = \mathbf{R}^T\mathbf{G}^T \qquad [6\text{-}8]$$

Because of [2-25] we know $\mathbf{R}$ has full rank. Thus from Lemma 4.1 we have the following theorem:

Theorem 6.1 .

A necessary condition for a matrix $\mathbf{E}$ to have a motion decomposition [6-4] is that $\mathbf{E}$ has exactly a rank two.

Proof: Because $\mathbf{R}$ has full rank thus the rank of $\mathbf{E}$ is equal to the rank of $\mathbf{G}$. Since $\mathbf{T} = [t_1\ t_2\ t_3]^T \neq 0$, thus a direct rank checking shows that the rank of $\mathbf{G}$ is exactly 2 in any cases. Q.E.D.

This theorem tells us that if $\mathbf{E}$ solved from [6-4] has a rank other than 2 then we cannot decompose it into $\mathbf{T}\times\mathbf{R}$. And if $\mathbf{E}$ can be decomposed into $\mathbf{T}\times\mathbf{R}$ then $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ must lie in a plane. In Theorem 6.3 we shall even show not every matrix of rank 2 can be decomposed into the form [6-2]. Again we have the following useful theorem:

Theorem 6.2

Let i, j and k be three incompatible indices among 1, 2, and 3, that is , $i\neq j$, $j\neq k$, $i\neq k$ and $i,j,k = \{1, 2, 3\}$, then in equation [6-8] $\mathbf{e}_i\times\mathbf{e}_j = 0$ iff $t_k = 0$, and $\mathbf{e}_i = 0$ iff $t_j = t_k = 0$ and $\mathbf{e}_j \neq 0$, $\mathbf{e}_k \neq 0$,$\mathbf{e}_j\times\mathbf{e}_k \neq 0$.

Proof: Because of property [2-28] we have

$$\mathbf{e}_i\times\mathbf{e}_j = \mathbf{R}^T(\mathbf{g}_i\times\mathbf{g}_j) \qquad [6\text{-}9]$$

From Lemma 5.1 we know $\mathbf{e}_i\times\mathbf{e}_j = 0$ iff $\mathbf{g}_i\times\mathbf{g}_j = 0$. But $\mathbf{g}_i\times\mathbf{g}_j = 0$ iff $t_k = 0$. Also from Lemma 5.1 we know that $\mathbf{e}_i = 0$ iff $\mathbf{g}_i = 0$ or iff $t_j = t_k = 0$, because $\mathbf{E}$ has rank two, thus the other two columns $\mathbf{e}_i$ and $\mathbf{e}_j$ of it cannot be zero at the same time as well as parallel to each other. So if $\mathbf{e}_i = 0$, then we must have $\mathbf{e}_j \neq 0$, and $\mathbf{e}_k \neq 0$,$\mathbf{e}_j\times\mathbf{e}_k \neq 0$. Q.E.D.

Noticing that $\mathbf{R}$ is an orthonormal matrix, we have

$$\mathbf{E}\ \mathbf{E}^T = \begin{bmatrix} \mathbf{e}_1^T \\ \mathbf{e}_2^T \\ \mathbf{e}_3^T \end{bmatrix} \begin{bmatrix} \mathbf{e}_1\,\mathbf{e}_2\,\mathbf{e}_3 \end{bmatrix} = \mathbf{G}\mathbf{R}\ \mathbf{R}^T\mathbf{G}^T = \mathbf{G}\mathbf{G}^T \qquad [6\text{-}10]$$

or we get

$$t_j^2 + t_k^2 = \|\mathbf{e}_i\|^2 \qquad [6\text{-}11]$$

$$t_i^2 = \frac{\|\mathbf{e}_j\|^2 + \|\mathbf{e}_k\|^2 - \|\mathbf{e}_i\|^2}{2} \qquad [6\text{-}12]$$

$$t_i\, t_j = -\mathbf{e}_i \cdot \mathbf{e}_j \qquad [6\text{-}13]$$

where i, j, k are three incompatible indices among 1, 2, and 3. Immediately we have

$$2(t_1^2 + t_2^2 + t_3^2) = 2\|\mathbf{T}\|^2 = \|\mathbf{e}_1\|^2 + \|\mathbf{e}_2\|^2 + \|\mathbf{e}_3\|^2 = \|\mathbf{h}\|^2 \qquad [6\text{-}14]$$

Thus if we scale $\mathbf{h}$ in advance such that $\|\mathbf{h}\|^2 = 2$, then we have $\|\mathbf{T}\| = 1$, that is, our solution of the translation is automatically a unit vector.

Now our algorithm becomes clear. First calculate $\|\mathbf{e}_m\|$, m = 1, 2, 3, let the smallest of them be $\|\mathbf{e}_i\|$. Then

$$t_i = \pm\sqrt{\frac{\|\mathbf{e}_j\|^2 + \|\mathbf{e}_k\|^2 - \|\mathbf{e}_i\|^2}{2}}$$

$$= \pm\sqrt{1 - \|\mathbf{e}_i\|^2}, \text{ when } \|\mathbf{h}\| = 1 \qquad [6\text{-}15]$$

and

$$t_j = -(\mathbf{e}_i \cdot \mathbf{e}_j) / t_i,\ j \neq i, j = 1,2,3. \qquad [6\text{-}16]$$

The absolute value of t_j and t_k can also be calculated from

$$t_j = \pm\sqrt{\frac{\|\mathbf{e}_i\|^2 + \|\mathbf{e}_k\|^2 - \|\mathbf{e}_j\|^2}{2}},\ j \neq k \neq i,\quad j, k = 1,2,3$$

$$= \pm\sqrt{1 - \|\mathbf{e}_j\|^2}, \text{ when } \|\mathbf{h}\| = 1 \qquad [6\text{-}16A]$$

[6-16] and [6-16A] should give the same value , otherwise we cannot find a consistent decomposition for the given $\mathbf{E}$. But in case there are some small errors in $\mathbf{E}$, then [6-16] and [6-16A] may not give the same results. So we suggest the following way to calculate t_m, m = 1,2,3. Using [6-15] and [6-16A] to decide the absolute value of t_m, m = 1,2,3, but using [6-16] to decide the signs of t_m, m =

1,2,3 for each set of $\mathbf{T}$. The reason is that a small change in $\mathbf{e}_i$ or $\mathbf{e}_j$ may cause a large change in $\mathbf{e}_i \cdot \mathbf{e}_j$ but can only cause a relatively smaller change in $\|\mathbf{e}_i\|^2 + \|\mathbf{e}_k\|^2 - \|\mathbf{e}_j\|^2$.

Thus, if $\mathbf{E}$ can be decomposed into $\mathbf{T} \times \mathbf{R}$, then we have two sets of solution of $\mathbf{T}$. For each set of $\mathbf{T}$ we have a corresponding solution for $\mathbf{R}$. Let the two sets of solutions be $\mathbf{T}_1, \mathbf{R}_1$ and $\mathbf{T}_2 = -\mathbf{T}_1, \mathbf{R}_2$.

Since $\|\mathbf{e}_i\|$ is the smallest of $\|\mathbf{e}_m\|$, m = 1, 2, 3, thus t_i cannot be zero and will be of the largest absolute value among t_m, m = 1,2,3. Notice that $\|\mathbf{g}_j\| = t_i^2 + t_k^2$, $j \neq i \neq k$, j,k = 1,2,3, and $\mathbf{g}_j \times \mathbf{g}_k = \pm t_i \mathbf{T}$, for any $i \neq j \neq k$, i,j,k = 1,2,3. Thus from Theorem 2.1 we know $\mathbf{R}$ can be recoverd from

$$\mathbf{R}^T = [\, \mathbf{e}_j \; \mathbf{e}_k \; \mathbf{e}_j \times \mathbf{e}_k \,][\, \mathbf{g}_j \; \mathbf{g}_k \; \mathbf{g}_j \times \mathbf{g}_k \,]^{-1} \qquad [6\text{-}17]$$

The inversion of right side of [6-17] is well conditioned because of our choosing of $\mathbf{g}_j$ and $\mathbf{g}_k$. By this way we find two sets of solutions for the decomposition of $\mathbf{E}$. Let the two sets of solutions be $\mathbf{T}_1, \mathbf{R}_1$ and $\mathbf{T}_2 = -\mathbf{T}_1, \mathbf{R}_2$. But only one set of solution is true. By the help of one correspondece we can decide the sign of $\mathbf{T}$ before we solve $\mathbf{R}$ and hence save some labor to find the true solution. Without losing generality, we assume the real solution of translation is $\mathbf{T} = \alpha \mathrm{T}_0$, where α is some constant and $\mathbf{T}_0$ is either $\mathbf{T}_1$ or $-\mathbf{T}_1$ and the real rotation is $\mathbf{R}_0$, where $\mathbf{R}_0$ could be either $\mathbf{R}_1$ or $\mathbf{R}_2$. Thus from [2-18] we have

$$\mathbf{Z}'\mathbf{\Theta}' = \mathbf{R}_0 \mathbf{\Theta} \mathbf{Z} + \alpha \mathbf{T}_0 \qquad [6\text{-}18]$$

Using $\mathbf{T}_0$ to cross-multiply both sides of [6-18] we get

$$\mathbf{T}_0 \times (\mathbf{Z}'\mathbf{\Theta}') = (\mathbf{T}_0 \times \mathbf{R}_0) \mathbf{\Theta} \mathbf{Z} \qquad [6\text{-}19]$$

or

$$\frac{\mathbf{Z}'}{\mathbf{Z}} \mathbf{G}_0 \mathbf{\Theta}' = \mathbf{E}\mathbf{\Theta} \qquad [6\text{-}20]$$

where G_0 is the matrix form of $\mathbf{T}_0 \times$ as in [6-1]. Now no matter $\mathbf{T}_0 = \mathbf{T}_1$ or $\mathbf{T}_0 = -\mathbf{T}_1$ $\frac{Z'}{Z}$ can be uniquely decided by

$$\frac{\mathbf{Z}'}{\mathbf{Z}} = \sqrt{\frac{\mathbf{\Theta}^T \mathbf{E}^T \mathbf{E} \mathbf{\Theta}}{\mathbf{\Theta}'^T \mathbf{G}_0^T \mathbf{G}_0 \mathbf{\Theta}'}} \qquad [6\text{-}21]$$

Since all other terms except $\mathbf{G}_0$ in [6-21] are fixed, thus in general if $\mathbf{T}_1$ satisfies [6-22], then $-\mathbf{T}_1$

will not, and vice versa. The only exception happens when $\mathbf{E\Theta} = 0$. In this case we cannot decide the true one. However, because $\mathbf{E}$ has rank 2, there will be at most one point (x,y) in the image plane such that $\mathbf{E\Theta} = 0$. To make the algorithm more robust we can modify the criterion as following. Let

$$L_1 = \sum_i \|\mathbf{E\Theta}_i - \frac{Z'_i}{Z_i}\mathbf{G}_1\mathbf{\Theta}'_i\|^2 \qquad [6\text{-}22]$$

$$L_2 = \sum_i \|\mathbf{E\Theta}_i + \frac{Z'_i}{Z_i}\mathbf{G}_1\mathbf{\Theta}'_i\|^2 \qquad [6\text{-}23]$$

where $\mathbf{G}_1$ is the matrix form of $\mathbf{T}_1$. Then if $L_1 < L_2$ then $\mathbf{T}_1$ and $\mathbf{R}_1$ is the solution. Otherwise we must have $L_2 < L_1$ and $\mathbf{T}_2$ and $\mathbf{R}_2$ is the solution.

So far we have decided $\mathbf{R}_0$ and $\mathbf{T}_0$, but the sign of α is still not decided. By requiring a positive depth and using [2-55] one can easily decide the sign of α.

In the following we shall give a necessary and sufficient condition under which a 3 by 3 matrix $\mathbf{E} = [\,\mathbf{e}_1\,\mathbf{e}_2\,\mathbf{e}_3\,]$ of rank 2 can be decomposed into $\mathbf{T}\times\mathbf{R}$, where $\mathbf{R}$ is a rotation matrix, and $\mathbf{T} \neq 0$ is an vector. We first list all the necessary conditions and then show that these are also sufficient. Notice [6-12] must be true for any three incompatible indices among 1, 2, and 3, thus we have the following necessary condition by requiring t_i, i = 1, 2,3, be real numbers:

$$\|\mathbf{e}_m\|^2 + \|\mathbf{e}_n\|^2 \geq \|\mathbf{e}_l\|^2, \quad l \neq m \neq n, \; l,m,n = 1,2,3 \qquad [6\text{-}24]$$

And [6-13] must be true for all $i \neq j$. Note that among the productions $t_1\,t_2$, $t_1\,t_3$ and $t_2\,t_3$, whatever the real numbers t_1, t_2, t_3 can be, we will never have two positive productions or three negative productions at the same time. Thus we must have:

neither exactly two of $\mathbf{e}_1\cdot\mathbf{e}_2$, $\mathbf{e}_1\cdot\mathbf{e}_3$, $\mathbf{e}_2\cdot\mathbf{e}_3$ are strictly negative

nor all of $\mathbf{e}_1\cdot\mathbf{e}_2$, $\mathbf{e}_1\cdot\mathbf{e}_3$, $\mathbf{e}_2\cdot\mathbf{e}_3$ are strictly positive [6-25]

Since Theorem 6.1 requires $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ be coplanar, thus [6-25] indicates that the three vectors can only have three configurations as shown in Fig. 6.1. Because [6-12] and [6-13] must be satisfied at the same time, then we must have

$$4(\mathbf{e}_l\cdot\mathbf{e}_m)^2 = (\|\mathbf{e}_m\|^2 + \|\mathbf{e}_n\|^2 - \|\mathbf{e}_l\|^2)(\|\mathbf{e}_l\|^2 + \|\mathbf{e}_n\|^2 - \|\mathbf{e}_m\|^2)$$
$$l \neq m \neq n, \; l,m,n = 1,2,3 \qquad [6\text{-}26]$$

Theorem 6.2 requires another necessary condition:

$$\text{if } \mathbf{e}_j \times \mathbf{e}_k \neq 0, \text{ then } t_i \neq 0, \text{ or } \|\mathbf{e}_j\|^2 + \|\mathbf{e}_k\|^2 \neq \|\mathbf{e}_i\|^2, i \neq j \neq k \tag{6-27}$$

After these conditions are satisfied then we can always solve t_i, and hence $\mathbf{g}_i$, i = 1,2,3, from [6-15] and/or [6-16]. And then we can solve $\mathbf{R}$ from [6-17]. Let's assume $\mathbf{e}_j$ and $\mathbf{e}_k$ are two independent vectors in $\mathbf{E}$. Then the third vector $\mathbf{e}_i$ must be a linear combination of $\mathbf{e}_j$ and $\mathbf{e}_k$. And $\mathbf{g}_i$ is also a linear combination of $\mathbf{g}_j$ and $\mathbf{g}_k$. Notice that our calculation of $\mathbf{R}$ from [6-17] automatically satisfies

$$\mathbf{e}_j = \mathbf{R}^T\mathbf{g}_j \text{ and } \mathbf{e}_k = \mathbf{R}^T\mathbf{g}_k, \tag{6-28}$$

so what we need to satisfy is $\mathbf{e}_i = \mathbf{R}^T\mathbf{g}_i$. Notice that from [6-1], [6-12] and [6-13] we have

$$\mathbf{g}_i = -\frac{t_j}{t_i}\mathbf{g}_j - \frac{t_k}{t_i}\mathbf{g}_k = \frac{2}{\|\mathbf{e}_j\|^2 + \|\mathbf{e}_k\|^2 - \|\mathbf{e}_i\|^2}((\mathbf{e}_j \cdot \mathbf{e}_i)\mathbf{g}_j + (\mathbf{e}_k \cdot \mathbf{e}_i)\mathbf{g}_k) \tag{6-29}$$

Thus by requiring $\mathbf{e}_i = \mathbf{R}^T\mathbf{g}_i$ and using [6-28], [6-29] we must have

$$\mathbf{e}_i = \frac{2}{\|\mathbf{e}_j\|^2 + \|\mathbf{e}_k\|^2 - \|\mathbf{e}_i\|^2}((\mathbf{e}_j \cdot \mathbf{e}_i)\mathbf{e}_j + (\mathbf{e}_k \cdot \mathbf{e}_i)\mathbf{e}_k) \tag{6-30}$$

Now let's derive an equivalent condition for [6-30]. Assume

$$\mathbf{e}_i = c_j\mathbf{e}_j + c_k\mathbf{e}_k = [\,\mathbf{e}_j \;\; \mathbf{e}_k\,]\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \tag{6-31}$$

Left-multiply both sides of [6-31] by $[\,\mathbf{e}_j\,\mathbf{e}_k\,]^T$ we get

$$\begin{bmatrix} \mathbf{e}_j^T\mathbf{e}_i \\ \mathbf{e}_k^T\mathbf{e}_i \end{bmatrix} = \begin{bmatrix} \mathbf{e}_j^T\mathbf{e}_j & \mathbf{e}_j^T\mathbf{e}_k \\ \mathbf{e}_k^T\mathbf{e}_j & \mathbf{e}_k^T\mathbf{e}_k \end{bmatrix}\begin{bmatrix} c_j \\ c_k \end{bmatrix} \tag{6-32}$$

Because $\mathbf{e}_j$ is not parallel to $\mathbf{e}_k$, then,

$$\Delta \triangleq \|\mathbf{e}_j\|^2\|\mathbf{e}_k\|^2 - (\mathbf{e}_j \cdot \mathbf{e}_k)^2 > 0 \tag{6-33}$$

Thus we can solve [6-32] to get

$$c_j = \frac{\|\mathbf{e}_k\|^2(\mathbf{e}_j \cdot \mathbf{e}_i) - (\mathbf{e}_j \cdot \mathbf{e}_k)(\mathbf{e}_k \cdot \mathbf{e}_i)}{\Delta}$$

$$c_k = \frac{\|\mathbf{e}_j\|^2(\mathbf{e}_k \cdot \mathbf{e}_i) - (\mathbf{e}_j \cdot \mathbf{e}_k)(\mathbf{e}_j \cdot \mathbf{e}_i)}{\Delta} \tag{6-34}$$

Because c_j and c_k must be unique, so comparing [6-30] with [6-31] we have

$$\frac{2\mathbf{e}_j \cdot \mathbf{e}_i}{\|\mathbf{e}_j\|^2 + \|\mathbf{e}_k\|^2 - \|\mathbf{e}_i\|^2} = \frac{\|\mathbf{e}_k\|^2(\mathbf{e}_j \cdot \mathbf{e}_i) - (\mathbf{e}_j \cdot \mathbf{e}_k)(\mathbf{e}_k \cdot \mathbf{e}_i)}{\|\mathbf{e}_j\|^2\|\mathbf{e}_k\|^2 - (\mathbf{e}_j \cdot \mathbf{e}_k)^2}$$

$$\frac{2\mathbf{e}_k\cdot\mathbf{e}_i}{\|\mathbf{e}_j\|^2+\|\mathbf{e}_k\|^2-\|\mathbf{e}_i\|^2} = \frac{\|\mathbf{e}_j\|^2(\mathbf{e}_k\cdot\mathbf{e}_i)-(\mathbf{e}_j\cdot\mathbf{e}_k)(\mathbf{e}_j\cdot\mathbf{e}_i)}{\|\mathbf{e}_j\|^2\|\mathbf{e}_k\|^2-(\mathbf{e}_j\cdot\mathbf{e}_k)^2} \qquad [6\text{-}35]$$

Finnally notice that if [6-24] ~ [6-27] and [6-35] are satisfied, then from our procedure listed above, we can always find two decompositions of **E** into the form $\mathbf{T}\times\mathbf{R}$. Thus in summary we have the following theorem:

Theorem 6.3 .

A matrix $\mathbf{E} = \begin{bmatrix} \mathbf{e}_1^T \\ \mathbf{e}_2^T \\ \mathbf{e}_3 \end{bmatrix}$ of rank 2 with $\mathbf{e}_j\times\mathbf{e}_k \neq 0$, $j \neq k$, $j,k = 1,2,3$, can be decomposed into the form $\mathbf{T}\times\mathbf{R}$, where **T** is a nonzero vector and **R** is a rotation matrix, iff [6-35] and [6-24] ~ [6-27] are satisfied. ||

Theorem 6.1 tells us only matrices of rank two can be decomposed into the form $\mathbf{T}\times\mathbf{R}$. Now Theorem 6.3 tells us not every matrix of rank two can find a decomposition into $\mathbf{T}\times\mathbf{R}$. But the least square solution of **E** from [6-4] cannot guarantee the conditions in Theorem 6.1 and Theorem 6.3 due to the noise effect. So intrinsically the unconstrained two step 8-point linear algorithm will not assure a consistent solution of **T** and **R**. Thus we argue, in practice, the unconstrained least square solution should be replaced by constrained least square solution and then we'll lose the linearity and the closedness of the solution at all. For this reason, we introduce our 3-point algorithm which is discussed in section 8.

One thing still in doubt is that whether a small violation in **E** against the conditions in Theorem 6.1 and 6.3 will cause a large variation in **T** and **R** if one insists to use some $\mathbf{T}\times\mathbf{R}$ to approximate **E** and which algorithm is the most robust one.

6.3. Recover R when T = 0

In case of pure rotation, the above algorithm may not give robust solution of **R** if noise exists, though sometimes it does give the approximate solution of **R**. We find our following algorithm is quite robust in case of pure rotation. Theorem 2.2 tells us that two point correpondences decide the rotation uniquely in case of pure rotation. Now our problem is how to integrate multiple correspondences to get a robust solution other than to use only two correspondences. Since the surface is general and the solution of rotation parameters is intrinsically nonlinear, so it seems difficult to find a least square solution. Inspired by the central limit theorem, we instead adopt the following solution.

Let (x_i, y_i) and (x'_i, y'_i), $i = 1,2, \cdots, n$, be the correspondence pairs. Divide the indices into two groups S_1 and S_2 according to the geometrical position of (x_i, y_i) such that the center of points in S_1 is distant from the center of those in S_2. And let γ_i be defined as [2-32]. Let

$$\delta_j = \sum_{i\in S_j} \Theta_i = \sum_{i\in S_j} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}, \quad j = 1,2 \qquad [6\text{-}36]$$

$$\delta'_j = \sum_{i\in S_j} \gamma_i \Theta'_i = \sum_{i\in S_j} \gamma_i \begin{bmatrix} x'_i \\ y'_i \\ 1 \end{bmatrix}, \quad j = 1,2 \qquad [6\text{-}37]$$

Then because of [2-31] we must have

$$\delta'_j = \mathbf{R}\,\delta_j, \quad j = 1,2 \qquad [6\text{-}38]$$

From Theorem 2.2 we can then recover $\mathbf{R}$ from

$$\mathbf{R} = [\,\delta'_1\ \delta'_2\ \delta'_1\times\delta'_2\,][\,\delta_1\ \delta_2\ \delta_1\times\delta_2\,]^{-1} \qquad [6\text{-}39]$$

Since we assume the center of S_1 is distant from that of S_2, the inversion in [6-39] is well defined. This calculation is supported by the central limit theorem. And we find even if there is a small translation (when we compare the relative significance of the translation and the rotation we compare the motion they caused in the image plane), [6-39] still gives good results only if S_1 and S_2 are appropriately selected.

There is still a problem here. How does one know the motion is a pure rotation? One way to test the pure rotation situation is following. First try to solve the rotation matrix $\mathbf{R}$ using any method. Then calculate the average distances between the rotated correspondences and the real correspondences, i.e., calculate

$$D_1 = \frac{1}{n}\sum_{i=1}^{n}((x'_i - x''_i)^2 + (y'_i - y''_i)^2) \qquad [6\text{-}40]$$

where

$$x''_i = \frac{r_{11}\, x_i + r_{12}\, y_i + r_{13}}{r_{31}\, x_i + r_{32}\, y_i + r_{33}} \qquad [6\text{-}41]$$

$$y''_i = \frac{r_{21}\, x_i + r_{22}\, y_i + r_{23}}{r_{31}\, x_i + r_{32}\, y_i + r_{33}}$$

And let the average motion in the image plane be

$$D_0 = \frac{1}{n}\sum_{i=1}^{n}((x'_i - x_i)^2 + (y'_i - y_i)^2) \qquad [6\text{-}42]$$

then, if

$$D_1 \ll D_0 \qquad [6\text{-}43]$$

the motion is likely a pure rotation, otherwise, either the sovled **R** is wrong, or the motion includes a translation.

7. Criteria for Optimal Motion Estimation

Before we propose our algorithm for general motion solution we need to discuss something about the performance criterion since our goal is to get a globally optimal solution. So far few criterions have been proposed to judge the best estimation of motion. The least square criterion has been used to give a linear solution of motion which is optimal only in that particular approach. In last section we have seen that the unconstrained least square method cannot guarantee a consistent motion solution and hence the solution of that method may not be meaningful, not to say the accuracy. However it still can be used to serve the performance judge for general motion estimation if used properly. Let's state it formally in the following.

First consider the case where $\mathbf{T} \neq 0$. Then given n pairs of correspondences the optimal solution should make

$$D_2 \triangleq \|\mathbf{A}\ \mathbf{h}\|^2 = \mathbf{h}^T(\mathbf{A}^T\mathbf{A})\,\mathbf{h} \qquad [7\text{-}1]$$

minimum, where **A** and **h** are defined by [6-2] and [6-5]. The advantage of this cirterion is again the linearity and simplicity in computation since the matrix $\mathbf{A}^T\mathbf{A}$ can be calculated in advance and does not change with the estimation of motion parameters $\mathbf{T}$ and $\mathbf{R}$. Note here we use [7-1] in a different way than we use it in section 6. In section 6 we use it to solve $\mathbf{h}$ and then to solve $\mathbf{T}$ and $\mathbf{R}$. But here we use it to give a performance measure for a given motion estimation. To use [7-1] properly one should first normalize the norm of translation vector $\mathbf{T}$ to the same level, e.g., let $\|\mathbf{T}\| = 1$.

The disadvantage of the above criterion is that each point plays a unequal role in t he performance decision depending on the geometrical positions of its correspondence and itself in the image plane. To make the algorithm more robust we wish the importance of each correspondence pair be equalized. Hence we introduce the following criterion. Rewrite the motion epipolar line constraint [2-38] as

$$[\, x' \; y' \; 1 \,][\, a(x,y) \; b(x,y) \; c(x,y) \,]^T = 0 \qquad [7\text{-}2]$$

where

$$[\, a(x,y) \; b(x,y) \; c(x,y) \,]^T = (\mathbf{T} \times \mathbf{R})[\, x \; y \; 1 \,]^T \qquad [7\text{-}3]$$

then we see, if motion estimation is perfect and there is no error in the correspondence then the correspondence (x', y') of any point (x, y) should lie on its motion epipolar line defined by[7-2]. Hence the average distance of correspondences to their motion epipolar lines gives a performance measure of the motion estimation. That is, our new criterion is

$$D_3 \overset{\Delta}{=} \sum_i d_i \qquad [7\text{-}4]$$

where

$$d_i = \frac{| a(x_i,y_i)x'_i + b(x_i,y_i)y'_i + c(x_i,y_i) |}{\sqrt{a^2(x_i,y_i) + b^2(x_i,y_i)}} \qquad [7\text{-}5]$$

is the distance of (x'_i,y'_i) to its estimated motion epipolar line. The computation of D_3 is more complex and takes much more time, but the roles of the correpondences are equalized. From our experiences D_3 works more robustly than D_2 as expected.

Now let's consider the pure rotation case. Since in this case $\mathbf{T} = 0$, D_2 and D_3 are of little significance. In stead, D_1 defined in section 6 should be used to judge the optimal estimation of rotation.

In general case, we may first compare D_2 or D_3 and then calculate D_1. If D_1 is very small, then D_1 serves the judge, otherwise D_2 or D_3 serves the judge.

In the pure rotation case, since no structure information can be got from motion, hence it seems to us D_1 would be the only criterion for optimality of rotation estimation. But if the motion involves a translation then many other structural constraints may be incorporated in to give more robust performance. D_2 and D_3 we discussed in this section are all based on the motion epipolar constraint and do not take into account the structure information contained in the correspondences such as rigidity and orientation constraints, which we discussed in section 3. Howver we would like to mention here that in case we can get many correspondences then the criteria we discuss here will generally suffice, but if we can only get few correspondences then, the structural information may be not only important for the robustness of the algorithm but critical for a unique motion solution. In fact since the relative depths can all be calculated from [2-54] after the motion is solved if the translation is not zero, then every motion constraint gives a credit to the motion estimation. So we can expect if more constraints are adopted in the performance criteria then the algorithm should be more robust. Much work is still needed to be done in this aspect.

8. 3-point Algorithm for General Motion Solution

In this section we introduce our algorithm for general motion solution from point correspondences. Our method is based on the plane motion solution discussed in section 5. The solution is linear, but we need to search in a reasonable space to find the globally optimal solution. Although we use D_1 ,D_2, and D_3 as the performance criteria, this algorithm can give the globally optimal solution for any performance criterion because we search. And given a number of correspondences, if the motion is uniquely defined by them (though we don't know the necessary and sufficient condition under which the motion is uniquely defined so far), then theoreticall our algorithm can always find the right solution if the criterion is approporiate. Now let's discuss our algorithm.

In section 2 we have shown if depths are known in both frames of images we may solve the motion uniquely with three correspondences. In section 5 we have shown if the depths are known in one frame of image, then three correspondences give up to four sets of motion solution. And the spurious solution can be figured out by applying criterion D_2 or D_3 or many other constraints. But in this paper, our attention is put on monocular vision, hence we do not assume we have any knowledge of the depth. Of course sequentially we can get the relative depth information from motion, but to avoid the error accumulation we have to have an initializing algorithm. One of our goal in motion analysis is to get the solution from as few as possible correspondences and another is to make the solution as robust as possible by adopting as much as possible information only if the computation time is reasonable. Since three correspondencs in image plane without knowing the depth cannot give a finite number of motion solutions thus we need at least four correspondences to get a finite number of solutions. In section 3 we have seen that four general surface correspondences will give a group of high dimensional polynomial equations of multiple variables and hence it seems a closed form solution is very difficult to get. But in section 5 we have seen four noncolinear but coplanar correspondences give up to two sets of motion solution in a closed form. That's why we use a planar model to solve general motion. The problem for this method is: how does one know which four points are coplanar in the space. And another problem is: how can one find a globally optimal solution satisfying all motion constraints given multiple correspondences? These problems are solved by searching.

Since the motion problem is intrinsically a nonlinear problem we understand that any linearization will introduce some intermediate variables and may violate some constraints in case of noise as we see in the 8-point linear algorithm. To solve a nonlinear problem without colsed form solution the only way is to search. The problem now becomes where to search? How large is the searching space? We find by the help of three correpondences the searching space of the motion parameters is confined by the possible positions of a fourth correspondence. We show this idea in Fig. 8-1. Assume P_i, $i = 1,2,3$, are three noncolinear points in image frame 1, and P'_i, $i = 1,2,3$ are their correspondences in image frame 2 respectively. We first trust these three correspondences. It is obvious that the three scene points corresponding to P_1, P_2, P_3 construct a plane in space. The same argument is true for P'_1, P'_2, P'_3. And the two planes are related by the motion. Let C_1 and C_2 be the central point of line $\overline{P_1P_2}$ and line $\overline{P_1P_3}$ respectively, and S be the intersection point of line $\overline{P_2C_2}$ and line $\overline{P_3C_1}$. We assume C_1, C_2 and S are all on the same plane that P_1, P_2 and P_3 define. These three points are just artificial and hence may not be seen in the image because the real scene surface may be curved. However we know that the correspondence of C_1 in image frame 2 must lie at some point between the line interval $\overline{P'_1P'_3}$ and the correspondence of S in the image frame 2 must lie within the triangle $\Delta P'_1P'_2P'_3$. So immediately we see the searching space is no larger than the triangle $\Delta P'_1P'_2P'_3$ since the real motion is definitely defined by the four coplanar correspondences besides a dual motion. To further reduce the searching space, we first introduce the following theorem.

Theorem 8.1 .

Assume $(X_i, Y_i, Z_i) \longleftrightarrow (X'_i, Y'_i, Z'_i)$, i = 1,2,3 and

P: $P_i = (X_i, Y_i, Z_i) \rightarrow p_i = (x_i, y_i)$, $P'_i = (X'_i, Y'_i, Z'_i) \rightarrow p'_i = (x'_i, y'_i)$, i = 1,2,3

and (X_i, Y_i, Z_i), i = 1,2,3, are three colinear points in the space with their projections on the image plane satisfying

$$x_3 = (x_1 + x_2)/2, \quad y_3 = (y_1 + y_2)/2 \qquad [8\text{-}1]$$

(see Fig. 8.2, where $p' = ((x_1+x_2)/2,(y_1+y_2)/2)$). Then

$$|x'_3-(x'_1+x'_2)/2| = \frac{|d-1|}{d+1}\,\frac{|x'_2-x'_1|}{2},$$

$$|y'_3-(y'_1+y'_2)/2| = \frac{|d-1|}{d+1}\,\frac{|y'_2-y'_1|}{2} \qquad [8\text{-}2]$$

where

$$d = \frac{Z_1Z'_2}{Z'_1Z_2} \qquad [8\text{-}3]$$

Proof: See Appendix B. Q.E.D..

This theorem tells us if we use the center point of $\overline{P'_1P'_2}$ to approximate C'_1, then the error can be calculated from [8-2]. Similar argument can be applied to C'_2. Note that in [8-2], the coordinates of two end points can be assumed known since we know P'_i, for i = 1,2,3. The only thing we don't know is d, which is a function the depths of two end point in both images. But by assuming a smooth motion and high enough frame rate we can always confine the range of d to a very safe extent. For example, we can assume

$$|d-1| < 0.2 \quad \text{or} \quad 0.8 < d < 1.2. \qquad [8\text{-}4]$$

Sequentially we can have a prediction of d. But initially we need to search a relatively larger window. Now it is clear that the searching window of S' is defined by the searching range of C'_1 and C'_2. This is shown in Fig. 8.3.

To allow small errors in the three correspondences P'_i, $i = 1,2,3$, one can still search them in the neighborhoods around each correspondence, as shown in Fig. 8.4. This searching space is reasonablly small if the given three correspondences are reliable. After we decide the seaching space the problem left is to find the optimal solution. We use D_3 and D_1 or D_2 and D_1 as the performance criterion to select out the best solution by the help of many other correspondences. To make our algorithm clear, we list the procedures in the following:

Step 1. Initialization: including deciding the searhing window by controling the parameter d in [8-2].

Step 2. Pick up three points P_1, P_2, P_3 and their correspondences and calculate the center point position of the triangle center S. To make the algorithm robust, the three points should separate as far as possible to make use of the full resolution of the camera. But the seaching space of S' will be consequently large, of course.

Step 3. Produce four would-be correspondences in the four searching windows and to get up to two sets of solutions of motion by using the plane model for each set of four correspondences.

Step 4. Calculate the performance of the current motion solutions by the help of many other correspondences and adopting D_3 (or D_2) and D_1. It is found the more correspondences are adopted, the more robust is the algorithm. But the computation is of course heavier.

Step 5. Refresh the best motion solution by comapring the performance function values of the current motion solutions to that of the best one solved earlier.

Step 6. Repeat Step 2 to 5 until all possibilities are done.

Now let's discuss a little bit about the characteristics of this algorithm. This algorithm is based on the idea as: now that a closed form solution of global optimality is difficult to get why don't we search a little bit? The problem is how and where to search. Now in this algorithm we transform the search of the motion parameters to the search of the correspondence of an artificial point and confine the searching space to a reasonalbe extent. This is quite different from the traditional direct solution method. Another character of this algorithm is that we introduce an optimality criterion to the motion solution. In fact this algorithm will work for any performance criterion. This algorithm makes use of the full resolution of the camera and all possible correspondences and promise a consistent, unified and robust solution for any kind of surface (except line) and motion. We should mention here that the searching can be done hierarchily and parallelly. However, there is also a disadvantage in this algorithm. That is, it relies too much on the three correspondences used for constructing the triangle. If any one of the three correspondences is very errorneous or all the three points are infinitely far away, then we may not get the globally optimal solution in the specified searching window. So one must be careful in selecting the three points.

9. Long Range Motion and Short Range Motion

In this section we discuss the long range motion and the short range motion. In the past, the attentions of most researchers are focused on short range motion. To make matching correspondences easy, small motion is an inevitable assumption in the low level processing. However small motion in the image plane may cause the motion estimation nonrobust. But in large motion case the correspondences can hardly be got directly. And short range motion parameters are also useful for navigation, robots, depth estimation etc.. We treat this problem in this section.

Our opinion is that one'd better accumulate the short range correspondences to get long range correspondences to get long range motion, and then decompose the long range motion into short range motion. This procedure makes the motion estimation well-conditioned. Let's show our procedure. Assume we have a sequence of images $I_0, I_1, I_2, \cdots, I_n$ taken at time $t_0, t_1, t_2, \cdots, t_n$, and assume there be a reasonable number of point correspondences between each image pair. The correspondences between I_0 and I_n may not be directly available if the motion between t_0 and t_n is large. However, correspondences between t_{i-1} and t_i, $i = 1,2,3, \cdots ,n$ can be directly available if the sampling rate is high enough regarding to the motion. Thus we can pass the information from

t_1 to t_n to get indirect correspondences between t_n and t_0. However in the passing process we should be aware that errors will not accumulate. One way or maybe the best way to guarantee this is to do matching between I_i and I_0 with the information got from the matching between I_{i-1} and I_0 and between I_i and I_{i-1}. After we get the matchings, we can now solve the long range motion. Let $\mathbf{R}_n, \mathbf{T}_n$ be the motion between t_n and t_0, for n = 1,2,3, $\cdots$, and let $\mathbf{R}_{n,n-1}, \mathbf{T}_{n,n-1}$ be the motion between t_n and t_{n-1}. Roughly we can say that $\mathbf{R}_n$ and $\mathbf{T}_n$ is long range motion and $\mathbf{R}_{n,n-1}$ and $\mathbf{T}_{n,n-1}$ is short range motion. Then from the correspondences between I_n and I_0 and between I_{n-1} and I_0, we can solve $\mathbf{R}_n, \mathbf{T}_n, \mathbf{R}_{n-1}$, and $\mathbf{T}_{n-1}$. From the motion relation we should have (see [5-72])

$$\mathbf{R}_n = \mathbf{R}_{n,n-1}\mathbf{R}_{n-1},$$
$$\mathbf{T}_n = \mathbf{T}_{n,n-1} + \mathbf{R}_{n,n-1}\mathbf{T}_{n-1} \qquad [9\text{-}1]$$

In stead of calculating the short range motion $\mathbf{R}_{n,n-1}, \mathbf{T}_{n,n-1}$ from the correspondences between I_n and I_{n-1}, we can calculate them from

$$\mathbf{R}_{n,n-1} = \mathbf{R}_n\mathbf{R}_{n-1}^{-1},$$
$$\mathbf{T}_{n,n-1} = \mathbf{T}_n - \mathbf{R}_{n,n-1}\mathbf{T}_{n-1} \qquad [9\text{-}2]$$

The reason for this methodology is easy to be seen. We all know that, to the extent the correspondences are available the larger the motion is, the more accurate the estimation of it will be. Even if direct matching between I_n and I_0 cannot be done, this method may still give better performance. Let's see it from a simple mathematical description. Assume the correspondence error ε in x (or y) is a white noise with variance σ^2. Thus in matching I_i and I_{i-1} we will have error ε_i, for i = 1,2,3, $\cdots$., and we assume the real motion in x is δx_i correspondingly. Because we pass the information from t_1 to t_n we will have the correspondence error between t_n and t_0 as

$$\eta = \sum_{i=1}^{n} \varepsilon_i \qquad [9\text{-}3]$$

and the real motion will be

$$\Delta x_n = \sum_{i=1}^{n} \delta x_i \qquad [9\text{-}4]$$

Obviously η is a zero mean noise of variance $n\sigma^2$. Because the motion is continuous in general, δx_i will add positively to each other most often. So we can expect the signal to noise ratio will have the relation

$$\frac{(\Delta x_n)^2}{n\sigma^2} = \frac{(\sum_{i=1}^{n} \delta x_i)^2}{n\sigma^2} \geq \frac{(\delta x_i)^2}{\sigma^2}, \text{ for most } i, \ 0 \leq i \leq n \qquad [9\text{-}5]$$

An numerical example is that $\delta x_2 = 2\delta x_1$, and we see $(\Delta x_2)^2 / 2\sigma^2 > (\delta x_i)^2 / \sigma^2$, for i = 1,2. This means [9-2] will be more robust in average than calculating $\mathbf{R}_{n,n-1}$ and $\mathbf{T}_{n,n-1}$ from the correspondences between I_n and I_{n-1}. If direct matching between I_n and I_0 is available, then ,the performance should be improved much more. Of course the description here is still too simple, but it supports the idea that short range motion estimation calculated from long range motion estimation may be more robust than that calculated directly from short range correspondences.

10. Experiment Results

The experiment results are to appear in the sebsequent papers for formal publication.

11. Summary

In this paper we presented some basic theorems, constraints and equations for the motion problem. These theoretical results could be essential for the motion estimation problem since they deal with some basic problems of motion and give simplest results. We also proposed several criteria for optimal motion estimation. We discussed several algorithms for general and planar motion solution. Almost all theorems in this paper have been directly or indirectly proved by experiments. And the algorithms introduced here seem superior to the existing similar algorithms if there is any. As long as no closed form nonlinear algorithm for general motion solution is available, our algorithm for general motion solution via planar model could be a good approach since it can minimize any given cost function by searching in a reasonable and predictable space.

In this paper we only deal with the perspective and monocular vision and we assume the correspondences are already available. In our following work we shall extend our results to other situations and work on the motion problem from the very beginning. The readers may also find that we are trying to unify the motion problem into a nice representation. Our results will be issued in a series of papers.

Appendix A

Assume $\mathbf{N}$ is solved and

$$\|\mathbf{N}\| = 1 \tag{A-1}$$

then from

$$\mathbf{R}^T\mathbf{R} = (\mathbf{K}-\mathbf{T}\mathbf{N}^T)^T(\mathbf{K}-\mathbf{T}\mathbf{N}^T) = \mathbf{I} \tag{A-2}$$

we have

$$\mathbf{K}^T\mathbf{K} - \mathbf{K}^T\mathbf{T}\mathbf{N}^T - \mathbf{N}\mathbf{T}^T\mathbf{K} + \mathbf{N}\mathbf{T}^T\mathbf{T}\mathbf{N}^T = \mathbf{I} \tag{A-3}$$

or

$$\|\mathbf{T}\|^2\mathbf{N}\mathbf{N}^T = \mathbf{I} - \mathbf{K}^T\mathbf{K} + \mathbf{K}^T\mathbf{T}\mathbf{N}^T + \mathbf{N}\mathbf{T}^T\mathbf{K} \tag{A-4}$$

Premultiply [A-4] by $\mathbf{N}^T$ and postmultiply it by $\mathbf{N}$ and use [A-1] we'll get

$$\|\mathbf{T}\|^2 = 1 - \|\mathbf{K}\mathbf{N}\|^2 + 2\mathbf{N}^T\mathbf{K}^T\mathbf{T} \tag{A-5}$$

Replace [A-5] into [A-4] and reorder it we have

$$(1-\|\mathbf{K}\mathbf{N}\|^2)\mathbf{N}\mathbf{N}^T + \mathbf{K}^T\mathbf{K} - \mathbf{I} = \mathbf{K}^T\mathbf{T}\mathbf{N}^T + \mathbf{N}\mathbf{T}^T\mathbf{K} - 2\mathbf{N}^T\mathbf{K}^T\mathbf{T}\mathbf{N}\mathbf{N}^T \tag{A-6}$$

Because of the symmetry [A-6] only gives us up to 6 different equations. So in the following we only list useful elements in matrices. Let

$$\mathbf{Q} = \begin{bmatrix} q_1 & q_2 & q_3 \\ q_2 & q_4 & q_5 \\ q_3 & q_5 & q_6 \end{bmatrix} = (1 - \|\mathbf{KN}\|^2)\mathbf{NN}^T + \mathbf{K}^T\mathbf{K} - \mathbf{I} \qquad \text{[A-7]}$$

$$\mathbf{K} = [\,\mathbf{k}_1\,\mathbf{k}_2\,\mathbf{k}_3\,] \qquad \text{[A-8]}$$

Then

$$\mathbf{K}^T\mathbf{TN}^T = \begin{bmatrix} \mathbf{k}_1^T\mathbf{T} \\ \mathbf{k}_2^T\mathbf{T} \\ \mathbf{k}_3^T\mathbf{T} \end{bmatrix} [\,n_1\,n_2\,n_3\,] = \begin{bmatrix} n_1\mathbf{k}_1^T\mathbf{T} & n_2\mathbf{k}_1^T\mathbf{T} & n_3\mathbf{k}_1^T\mathbf{T} \\ n_1\mathbf{k}_2^T\mathbf{T} & n_2\mathbf{k}_2^T\mathbf{T} & n_3\mathbf{k}_2^T\mathbf{T} \\ n_1\mathbf{k}_3^T\mathbf{T} & n_2\mathbf{k}_3^T\mathbf{T} & n_3\mathbf{k}_3^T\mathbf{T} \end{bmatrix} \qquad \text{[A-9]}$$

$$\mathbf{K}^T\mathbf{TN}^T + \mathbf{N}^T\mathbf{T}^T\mathbf{K} = \mathbf{K}^T\mathbf{TN}^T + (\mathbf{K}^T\mathbf{TN}^T)^T$$

$$= \begin{bmatrix} 2n_1\mathbf{k}_1^T\mathbf{T} & (n_1\mathbf{k}_2^T + n_2\mathbf{k}_1^T)\mathbf{T} & (n_1\mathbf{k}_3^T + n_3\mathbf{k}_1^T)\mathbf{T} \\ & 2n_2\mathbf{k}_2^T\mathbf{T} & (n_2\mathbf{k}_3^T + n_3\mathbf{k}_2^T)\mathbf{T} \\ & & 2n_3\mathbf{k}_3^T\mathbf{T} \end{bmatrix} \qquad \text{[A-10]}$$

$$-2\mathbf{N}^T\mathbf{K}^T\mathbf{TNN}^T = -2(n_1\mathbf{k}_1^T + n_2\mathbf{k}_2^T + n_3\mathbf{k}_3^T)\mathbf{T} \begin{bmatrix} n_1^2 & n_1n_2 & n_1n_3 \\ & n_2^2 & n_2n_3 \\ & & n_3^2 \end{bmatrix} \qquad \text{[A-11]}$$

Thus from [A-6] we have

$$\begin{aligned} q_1 &= (-2n_1^3+2n_1)\mathbf{k}_1^T + (-2n_1^2n_2)\mathbf{k}_2^T + (-2n_1^2n_3)\mathbf{k}_3^T\mathbf{T} \\ q_2 &= (-2n_1^2n_2+n_2)\mathbf{k}_1^T + (-2n_1n_2^2+n_1)\mathbf{k}_2^T + (-2n_1n_2n_3)\mathbf{k}_3^T\mathbf{T} \\ q_3 &= (-2n_1^2n_3+n_3)\mathbf{k}_1^T + (-2n_1n_2n_3)\mathbf{k}_2^T + (-2n_1n_3^2+n_1)\mathbf{k}_3^T\mathbf{T} \qquad \text{[A-12]} \\ q_4 &= (-2n_2^2n_1)\mathbf{k}_1^T + (-2n_2^3+2n_2)\mathbf{k}_2^T + (-2n_2^2n_3)\mathbf{k}_3^T\mathbf{T} \end{aligned}$$

$$q_5 = (-2n_1n_2n_3)\mathbf{k}_1^T + (-2n_3n_2^2+n_3)\mathbf{k}_2^T + (-2n_2n_3^2+n_2)\mathbf{k}_3^T\mathbf{T}$$

$$q_6 = (-2n_1n_3^2)\mathbf{k}_1^T + (-2n_2n_3^2)\mathbf{k}_2^T + (-2n_3^3+2n_3)\mathbf{k}_3^T\mathbf{T}$$

Finally rearrange [A-12] into a matrix form we get the equation [5-38]. A direct rank check procedure shows that the rank of $\mathbf{W}$ is always 3 if [A-1] holds.

Appendix B

The line l passing (X_1, Y_1, Z_1) and (X_2, Y_2, Z_2) can be represented by

$$X = S(X_2 - X_1) + X_1$$
$$Y = S(Y_2 - Y_1) + Y_1$$
$$Z = S(Z_2 - Z_1) + Z_1 \qquad [B-1]$$

where S is a parameter. Now assume (X_3, Y_3, Z_3) is a point on l such that

$$x_3 \triangleq \frac{X_3}{Z_3} = \frac{1}{2}\left(\frac{X_1}{Z_1} + \frac{X_2}{Z_2}\right) \triangleq \frac{(x_1 + x_2)}{2} \qquad [B-2]$$

then we must have some s such that

$$X_3 = s(X_2 - X_1) + X_1$$
$$Y_3 = s(Y_2 - Y_1) + Y_1$$
$$Z_3 = s(Z_2 - Z_1) + Z_1 \qquad [B-3]$$

and

$$\frac{s(X_2 - X_1) + X_1}{s(Z_2 - Z_1) + Z_1} = \frac{1}{2}\left(\frac{X_1}{Z_1} + \frac{X_2}{Z_2}\right) = x_3 \qquad [B-4]$$

Using [B-2] we can solve [B-4] to get

$$s = \frac{x_3Z_1 - X_1}{X_2 - X_1 - x_3(Z_2 - Z_1)} = \frac{\frac{X_2}{Z_2}Z_1 - X_1}{X_2 - X_1 - \left(\frac{X_1}{Z_1}Z_2 - \frac{X_2}{Z_2}Z_1\right)} \qquad [B-5]$$

It is easy to show that

$$\frac{1 - s}{s} = \frac{Z_2}{Z_1} \qquad [B-6]$$

Now assume the motion make $(X_i, Y_i, Z_i) \longleftrightarrow (X'_i, Y'_i, Z'_i)$, $i = 1,2,3$. Because of the rigidity condition we must also have

$$\begin{aligned} X'_3 &= s(X'_2 - X'_1) + X'_1 = sX'_2 + (1-s)X'_1 \\ Y'_3 &= s(Y'_2 - Y'_1) + Y'_1 = sY'_2 + (1-s)Y'_1 \\ Z'_3 &= s(Z'_2 - Z'_1) + Z'_1 = sZ'_2 + (1-s)Z'_1 \end{aligned} \qquad [B-7]$$

To prove [B-7] one just need require

$$(Z'_j - Z'_1)^2 + (Y'_j - Y'_1)^2 + (X'_j - X'_1)^2 = (Z_j - Z_1)^2 + (Y_j - Y_1)^2 + (X_j - X_1)^2 \ , \ j = 2,3. \qquad [B-8]$$

and that (X_3, Y_3, Z_3) is inside (X_1, Y_1, Z_1) and (X_2, Y_2, Z_2). Now assume

$$\mathbf{P}: \ (X'_i, Y'_i, Z'_i) \rightarrow (x'_i, y'_i), \ i = 1,2,3$$

then we have

$$\begin{aligned} & x'_3 - (x'_1 + x'_2)/2 \\ &= \frac{sX'_2 + (1-s)X'_1}{sZ'_2 + (1-s)Z'_1} - \frac{1}{2}\left(\frac{X'_1}{Z'_1} + \frac{X'_2}{Z'_2}\right) \\ &= \frac{1}{2} \frac{2Z'_1Z'_2sX'_2 + 2(1-s)X'_1Z'_1Z'_2 - X'_2Z'_1[sZ'_2 + (1-s)Z'_1] - X'_1Z'_2[sZ'_2 + (1-s)Z'_1]}{Z'_2Z'_1[sZ'_2 + (1-s)Z'_1]} \\ &= \frac{1}{2} \frac{sZ'_1Z'_2X'_2 + (1-s)X'_1Z'_1Z'_2 - (1-s)X'_2Z'^2_1 - sX'_1Z'^2_2}{Z'_2Z'_1[sZ'_2 + (1-s)Z'_1]} \\ &= \frac{1}{2} \frac{(X'_2Z'_1 - X'_1Z'_2)(sZ'_2 - (1-s)Z'_1)}{Z'_2Z'_1[sZ'_2 + (1-s)Z'_1]} \\ &= \frac{\left(\frac{X'_2}{Z'_2} - \frac{X'_1}{Z'_1}\right)}{2} \frac{\frac{Z'_2}{Z'_1} - \frac{(1-s)}{s}}{\frac{Z'_2}{Z'_1} + \frac{(1-s)}{s}} \end{aligned} \qquad [B-9]$$

Using [B-6], we immediately have

$$|x'_3 - \frac{(x'_1 + x'_2)}{2}| = \frac{|d-1|}{d+1} \frac{|x'_2 - x'_1|}{2} \qquad [B-10]$$

where

$$d = \frac{Z_1Z'_2}{Z'_1Z_2} \qquad [B-11]$$

The same reason, we have

$$|y'_3 - \frac{(y'_1 + y'_2)}{2}| = \frac{|d-1|}{d+1} \frac{|y'_2 - y'_1|}{2} \qquad [B-12]$$

REFERENCES

1. Alessandro Verri and Tomaso Poggio, "Against Quantitative Optical Flow", ICCV-87, p. 171.

2. Ken-ichi Kanatani, "Structure and Motion from Optical Flow under Orthographic Projection", Computer Vision, Graphics and Image Processing 35, 181-199 (1986).

3. Ken-ichi Kanatani, "Structure and Motion from Optical Flow under Perspective Projection", Computer Vision, Graphics and Image Processing 38, 122-146 (1987).

4. Leo Levi, Applied Optics, Chap. 1, Vol. 1; Chap. 12, Vol. 2, John Wiley & Sons, 1980, 1967.

5. R.S. Longhurst, Geometrical and Physical Optics, 1973.

6. E.P. Lavin, Specular Reflection.

7. B.K.P, Horn and B.G. Schunk, "Determining Optical Flow", Artif. Intell. 17, 1981, 185-203.

8. K. Kanatani, "Transformation of Optical Flow by Camera Rotation", in Proceedings, IEEE Workshop on Motion: Representation and Analysis, Charleston, SC, 1986, pp. 111-118.

9. A.A. Giordano and F.M. Hsu, Least Square Estimation with Applications to Digital Signal Processing, John Wiley & Sons, 1985.

10. G.R. Fowler, Analytical Mechanics, Holt, Rinehart and Winston.

11. Stevens, and Yuille, "Depth Reconstruction in Stereopsis", ICCV-87, p. 682.

12. William Feller, An Introduction to Probability Theory and Its Applications, John Wiley & Sons, 1966.

13. A. Papoulis, Probability Random Variables and Stochestic Process, McGraw-Hill, New York, 1969.

14. R.M. Haralick and X. Zhuang, "A Note on 'Rigid Body Motion from Depth and Optical Flow'", Computer Vision, Graphics, and Image Processing 34, 373-387 (1986).

15. D.H. Ballard, and O.A. Kimball, "Rigid Body Motion from Depth and Optical Flow", Computer Vision Graphics and Image Processing 22, 1983, 95-115.

16. X. Zhuang and R.M. Haralick, "Rigid Body Motion and the Optical Flow Image", in the First Conf. on Artificial Intelligence Appl., Denver, Dec. 6-7, 1986.

17. R.Y. Tsai and T.S. Huang, "Estimating Three-Dimensional Motion Parameters of a Rigid Planar Patch", IEEE Transactions on Accoustics, Speech and Signal Processing, Vol. ASSP-29, No. 6, 1981.

18. R.Y. Tsai and T.S. Huang, "Estimating Three-Dimensional Motion Parameters of a Rigid Planar Patch, II: Singular Value Decomposition", IEEE Transactions on Accoustics, Speech and Signal Processing, Vol. ASSP-30, No. 4, 1982.

19. R.Y. Tsai and T.S. Huang, "Estimating Three-Dimensional Motion Parameters of a Rigid Planar Patch, III: Finite Point Correspondences and the Three-View Problem", IEEE Transactions on Accoustics, Speech and Signal Processing, Vol. ASSP-32, No. 2, 1984.

20. J.-Q. Fang, and T.S. Huang, "Solving Three-Dimensional Small-Rotation Motion Equations: uniqueness, Algorithms, and Numerical Results", Computer Vision, Graphics, and Image Processing 26, 183-206 (1984).

21. T.S. Huang, "Determining Three-Dimensional Motion and Structure from two Perspective Views", Chapter 14, Handbook of Pattern Recognition and Image Processing, Young & Fu (editors).

22. X. Zhuang, T.S. Huang, and R.M. Haralick, "Two-View Motion Analysis: A Unified Algorithm", Journal of Optical Society of America, A., vol. 3, No. 9, pp. 1492-1500, Sept., 1986.

23. C. Longuet-Higgins, "A Computer Algorithm for Reconstructuring A Scene from Two Projections", Nature 293 (Sept. 1981), pp.133-135.

24. C. Longuet-Higgins, "The Reconstruction of a scene from two projections -- Configurations that

defeat 8-point Algorithm", The First Conf. on A.I. Appl., Denver, Dec. 5-7, 1984.

25. J. Weng, N. Ahuja and T.S. Huang, Motion and Structure from Point Correspondences: A Robust Algorithm for Planar Case with Error Estimation, Proc. Inter. Conf. Pattern Recognition, Rome, Italy, 1988.

26. R.M. Winger, An Introduction to Projective Geometry. New York: Dover, 1962.

27. J. Weng, T.S. Huang and N. Ahuja, 3-D Motion Estimation, Understanding and Prediction from Noise image Sequences, IEEE Trans. Patern Anal. Machine Intell., Vol. PAMI-9, No. 3, pp.370-389, 1987.

28. C.H. Lee, "Structure and Motion from Two Perspective Views via Planar Patch", ICCV-88, Dec., 1988, Florida, pp. 158-164.

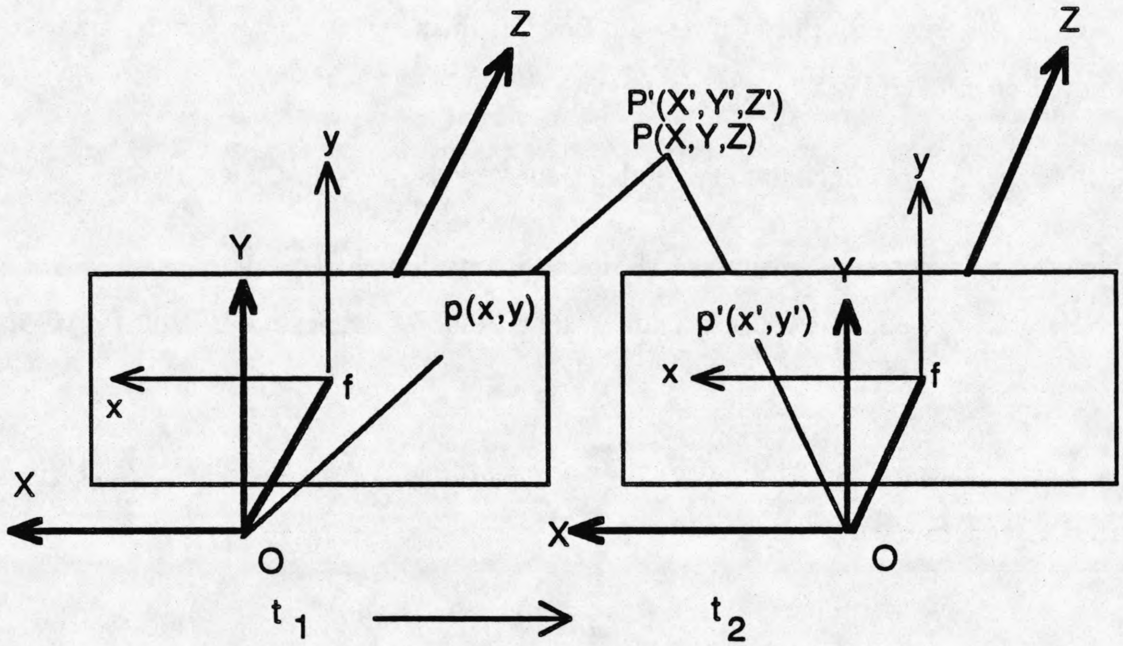


Fig. 2-1

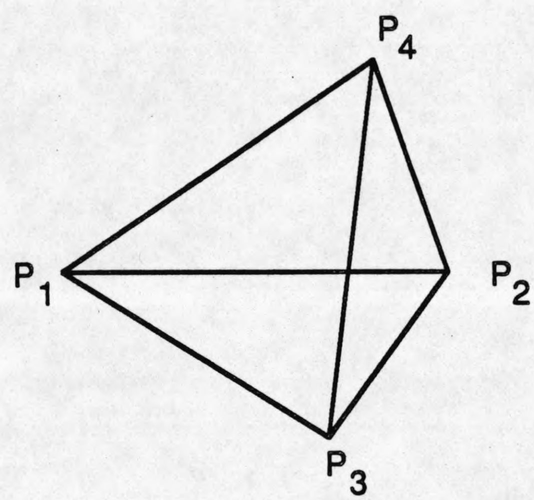


Fig. 3-1

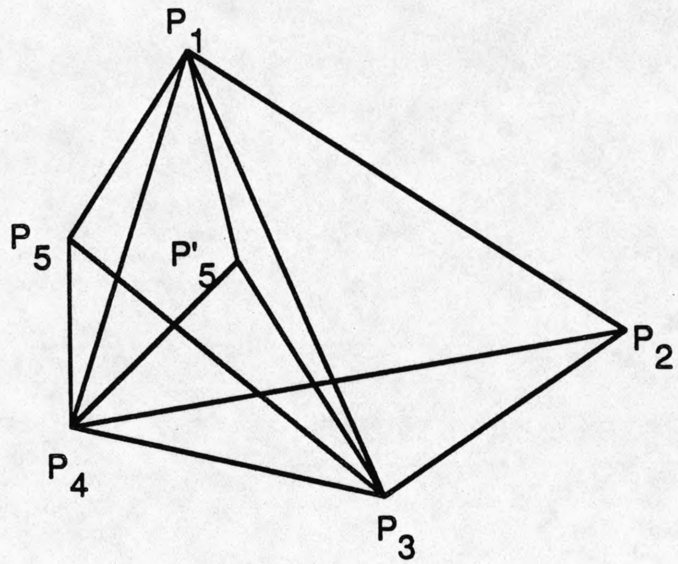

P1
P5
P'5
P2
P4
P3

Fig. 3-2

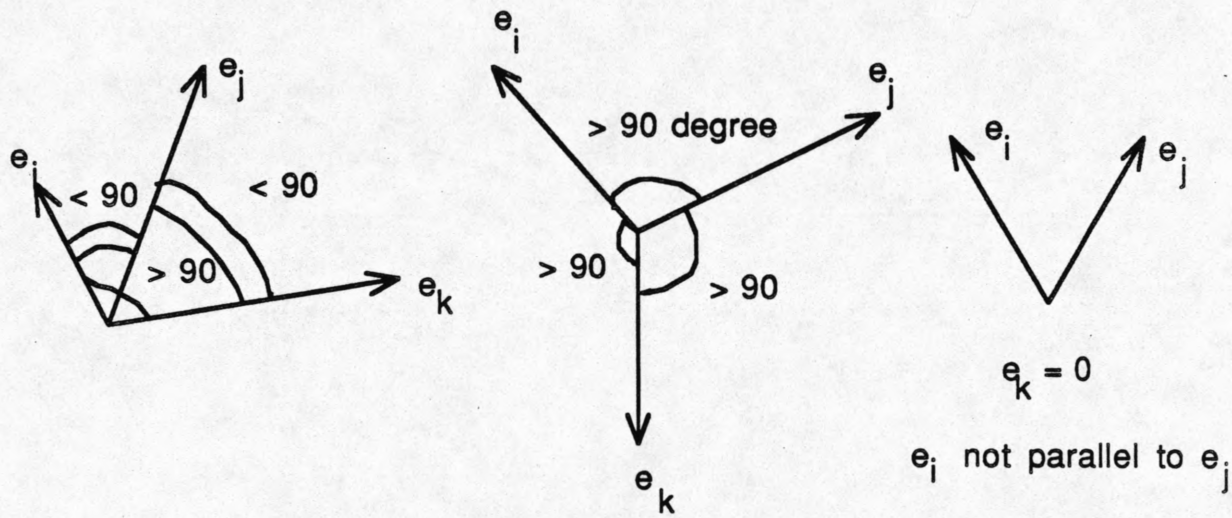
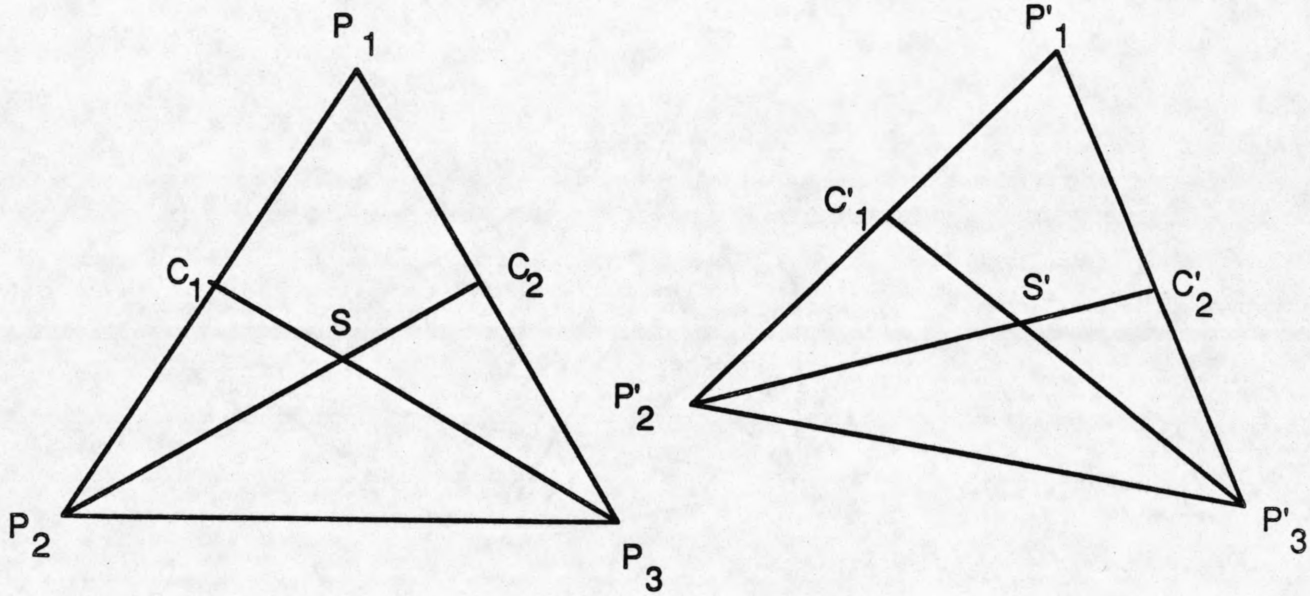
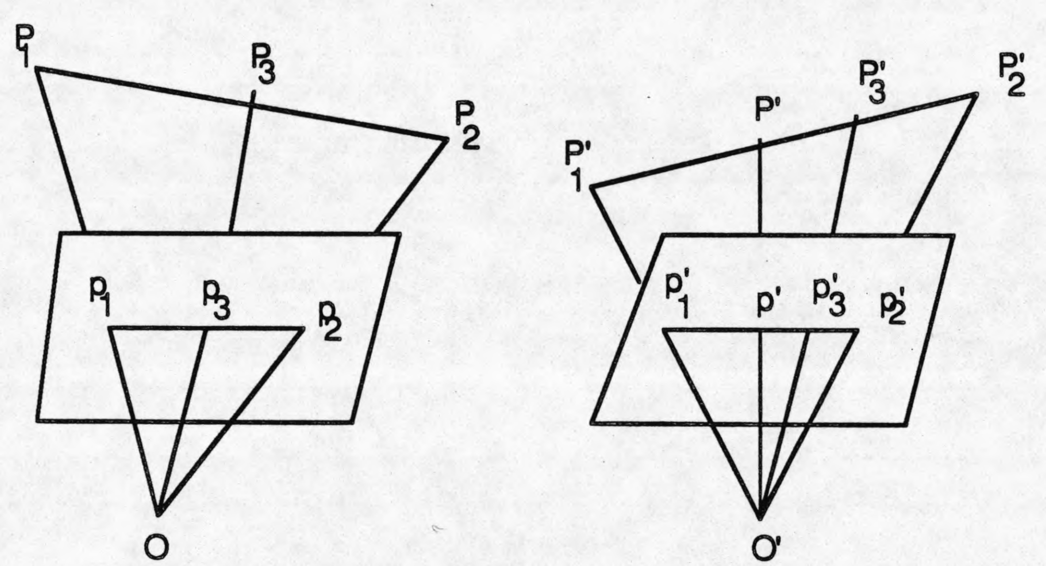

ej
ei
< 90
< 90
> 90
ek
ei
ej
> 90 degree
> 90
> 90
ek
ei
ej
ek = 0
ei not parallel to ej

Fig. 6.1



P_1
C_1
C_2
S
P_2
P_3
P'_1
C'_1
S'
C'_2
P'_2
P'_3

Fig. 8-1



P_1
P_3
P_2
p_1
p_3
p_2
O
P'_1
P'
P'_3
P'_2
p'_1
p'
p'_3
p'_2
O'

Fig. 8-2

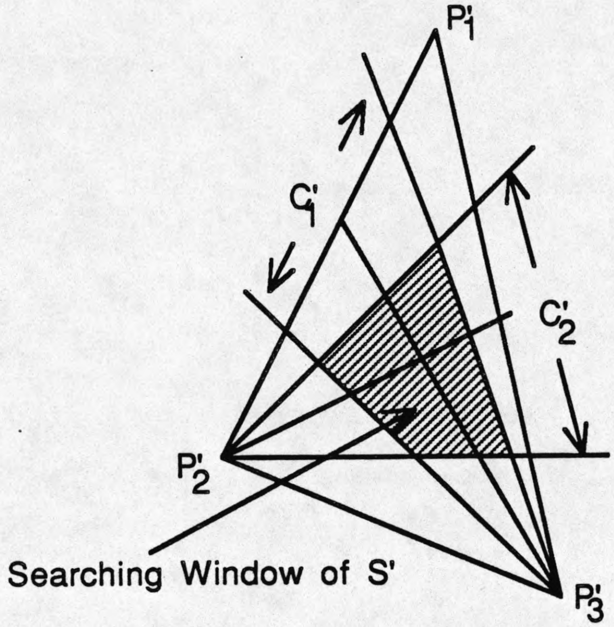

P'1
C'1
C'2
P'2
Searching Window of S'
P'3

Fig. 8-3

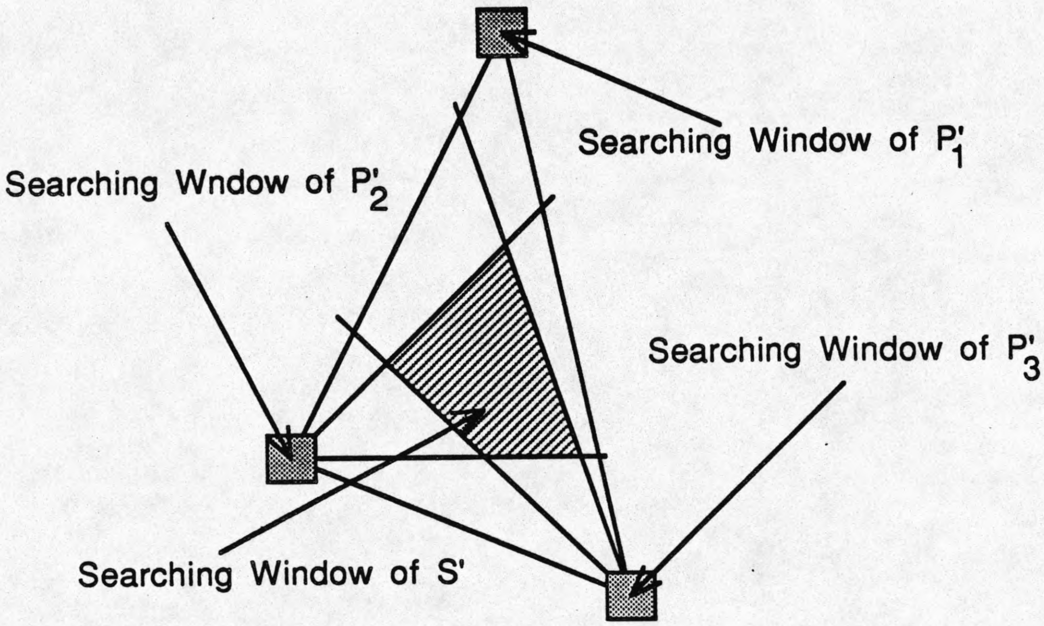

Searching Window of P'1
Searching Wndow of P'2
Searching Window of P'3
Searching Window of S'

Fig. 8-4