Chapter 8

Relative Orientation of Stereo Images

8.1 Introduction

Recovering the relative geometry of two stereo images is a classical problem in photogrammetry and also a basic problem in stereo vision. Technically, this problem includes interior orientation of each image and relative orientation of the two images, which are the prerequisite for object localization and surface reconstruction from the two images. Interior orientation of an image refers to the determination of 3 intrinsic parameters: the principal distance - also called focal length in an inexact sense - and the principal point position in a given image plane coordinate system. Relative orientation of two images refers to determination of the relative baseline (translation in motion vision) vector of two perspective centres and the relative rotation (matrix or angles) of one image relative to the other, which totally involves only five parameters. Therefore, the relative geometry of two stereo images includes totally 11 parameters (two sets of three intrinsic parameters and one set of five relative orientation parameters). However, if only homologous image point coordinates are measured, it is known that only seven parameters can be solved from the image measurements. Considering general application, the principal point position can be determined by using fiducial marks with metric cameras, or good cameras are manufactured such that the principal point can be very close to the image coordinate centre. Therefore, in the following discussions, we assume either the two principal points to be known *a priori* or sufficiently close to zero. Hence, we choose to study the problem of solving for two principal distance and five relative orientation parameters, a problem we shall call *general relative orientation.* This problem is chosen because it contains a selection of the maximum number of parameters which is symmetric and practically relevant. The word 'general' means that we intend to generalize the standard relative orientation problem in photogrammetry by including also two principal distances. The importance of this problem is obvious: given two stereo images with unknown interior and relative orientations, it is impossible to reconstruct 3D object points or surfaces without solving the general relative orientation even a general image matching can be achieved [Pan, 1996]. This paper presents a direct closed-form solution to the general relative orientation problem. We shall concentrate on the existence and correctness of this solution. The error sensitivity analysis is beyond the scope of this paper. However, a numerical example and a real image application are provided.

Camera calibration in photogrammetry terminology refers to determining the three intrinsic geometric parameters and all other camera distortions. This is routinely done by using control points. Self-calibration of cameras in photogrammetry refers to include the camera geometry and distortion parameters into the simultaneous adjustment of photogrammetric network. However, in computer vision literature, camera calibration refers to anything related to solving intrinsic and extrinsic geometry and distortion parameters of one, or two, or three cameras, thus may include interior and relative orientation. In a more general setup, we may have a network of overlapping images covering the object surfaces of interest completely. The recovery of the image overlapping topology and geometry of this image network may be better called *image resituation* [Pan et al, 1995a, 1995b], meaning to recover the topological and geometric situation of all images of this network. Therefore, image resituation may be understood as a general umbrella covering interior and relative orientation, and camera calibration of any number of networked images. However, in this paper, we shall concentrate on the general relative orientation of two stereo views, though this background is clarified.

Most photogrammetrists consider the problems of interior and relative orientation problem to be already solved. These problems have been revisited in the recently years by computer vision specialists partly because the early photogrammetric literature is not easily available to computer vision community, and largely because the stereo vision system, e.g. robot head, touches somes aspects of the stereo geometry which may be rather different from the standard aerial photogrammetric applications. It is also the case in close-range photogrammetry. The formulation of the stereo geometry (interior and relative orientation, etc.) goes back to the period around the turn of the century to 30's [Heisse, 1863; Finsterwalder, 1899, 1932; Fourcade, 1903; Kruppa, 1913]. Analytical photogrammetry including interior orientation, camera calibration, relative orientation, absolute orientation, up to aerial triangulation network adjustment were intensively studied from the 50's to the 70's [Schut, 1955; Thompson, 1959; Schmid, 1954; Abdel, 1971; Stefanovic, 1973; Wang, 1979; Khlebnikova, 1983]. The relative orientation problem was studied by computer vision community rather from a setup of motion vision [Longuet et al, 1980; Longuet, 1981; Bruss et al, 1983]. Horn (1990) bridged over the photogrammetry and computer vision communities, and gave a thorough description of relative orientation problem. In recent years, the problem of interior and relative orientation were revisited by many computer vision specialists as well as some photogrammetrists [Huang et al, 1989; Brandstätter, 1991, 1996; Faugeras et al, 1992; Hartley, 1992; Faugeras, 1993; Niini, 1993; Weng et al, 1993; Niini, 1994].

In the standard photogrammetric applications, relative orientation is solved through an iterative least-squares algorithm. A closed-form solution which is usually called *relative linear transform* was first formulated by Thompson (1959) and further exploited by Stefanovic (1973). As Horn (1990) pointed out, Longuet-Higgin's *essential matrix* [Longuet, 1981] is similar to Thompson's transform. Of course, a number of properties of Longuet-Higgin's *essential matrix* and Faugeras' *fundamental matrix* [Faugeras et al, 1992] discovered in recent years are not explicitly known previously. On the other hand, Faugeras' fundamental matrix has other usefulness in image matching and motion vision. The work presented here may be considered as a follow-up of these earlier works. We present a new direct closed-form approach to the general relative orientation problem, i.e. solving for two principal distances and five relative orientation parameters from pure image measurements. This problem was also tackled by Hartley (1992), but in a quite

different way which is rather more complicated and less direct than our approach.
In the next section, we first clarify the relative geometry of two stereo images. Central to the stereo geometry is the epipolar constraint which is expressed by the *coplanarity equation* for each pair of homologous image points. It is also known that in photogrammetric network adjustment, such as bundle adjustment, the *collinearity equations* characterizing the perspective projection of each single image play the central role, so the relative orientation of multiple images may be totally avoided, but at the expense of having control points or other information. However, the *coplanarity equations* represent still the basic stereo geometry if only two images are available. It is especially useful for robot stereo vision systems and biological stereo vision study.

8.2 Explicit and Implicit Coplanarity Equations

For the generality of stereo geometry (Fig. 8.1), let us assume that there is a 3D global coordinate system $O - XYZ$ with origin O and orthogonal axes labelled XYZ, within which two stereo images, left and right, are situated. The 2D image coordinate system $o - xy$ and the 3D camera coordinate system $C - \tilde{x}\tilde{y}\tilde{z}$ will be local to the left image, where C is the perspective centre of the left camera. Symbols relating to the right image will be marked with a prime $'$, yielding the right image coordinate systems $o' - x'y'$ and right camera coordinate system $C' - \tilde{x}'\tilde{y}'\tilde{z}'$. Let c denote the principal point [i.e. the orthogonal projection of the perspective centre on the image plane] of the left image and f the left principal distance, i.e. $f = \overline{Cc}$. The position C in $O - XYZ$ is denoted by (X_C, Y_C, Z_C), the position c in $o - xy$ is (x_c, y_c). The orientation of $C - \tilde{x}\tilde{y}\tilde{z}$ in $O - XYZ$ is captured by an orthonormal rotation matrix R defined by 3 successive rotation angles α, β, γ around $C - \tilde{x}$, $C - \tilde{y}$, and $C - \tilde{z}$ axes respectively.

8.2.1 Explicit Coplanarity Equation

When the only available information is a number of homologous image points, we can easily see that the collinearity equations given in (2.13) - (2.14) contain too many unknowns that cannot be solved from pure image measurements. Let $\mathbf{B}$ denote the baseline vector,

$$\mathbf{B} = \begin{pmatrix} B_x \\ B_y \\ B_z \end{pmatrix} = \begin{pmatrix} X_C - X_{C'} \\ Y_C - Y_{C'} \\ Z_C - Z_{C'} \end{pmatrix} \tag{8.1}$$

As the magnitude of $\mathbf{B}$ cannot be determined without control information, let $\mathbf{b}$ denote a positively scaled version of $\mathbf{B}$

$$\mathbf{b} = (b_x \ b_y \ b_z)^t = \kappa \mathbf{B}, \qquad \kappa > 0 \tag{8.2}$$

For any given object point P visible from the left and right images, there are two homologous image points $p(x, y)$ and $p'(x', y')$. A simple fact is that the five points P, p, p', C, and C' are coplanar, which can be captured by a cross product of three vectors being equal to zero

$$[\bar{\mathbf{p}}\,\mathbf{b}\,\bar{\mathbf{p}}'] = \bar{\mathbf{p}} \cdot (\mathbf{b} \times \bar{\mathbf{p}}') = 0 \tag{8.3}$$

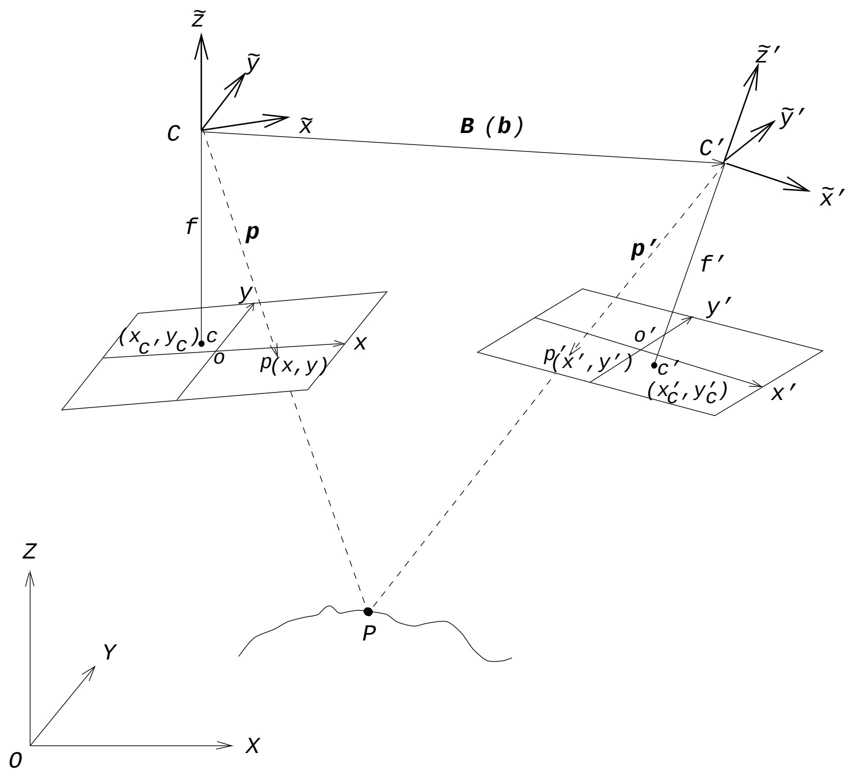

Figure 8.1: Analytical Geometry of Two Stereo Images

where

$$\mathbf{p} = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}, \quad \tilde{\mathbf{p}} = \Omega \mathbf{p}, \quad \bar{\mathbf{p}} = R\tilde{\mathbf{p}} \tag{8.4}$$

and Ω is defined as in (2.8), representing the translation of the image coordinate system origin to the principal point (x_c, y_c, f).
This is the coplanarity equation so-called traditionally in photogrammetry. There are several ways of rewriting this coplanarity equation. One way is to write it in the simplest analytical form

$$\begin{vmatrix} b_x & b_y & b_z \\ \bar{x} & \bar{y} & \bar{z} \\ \bar{x}' & \bar{y}' & \bar{z}' \end{vmatrix} = 0 \tag{8.5}$$

This is the most direct form of coplanarity equation used in iterative relative orientation in photogrammetry.
We can also rewrite the coplanarity equation (8.3) in the form of matrix product,

$$\bar{\mathbf{p}} \cdot (\mathbf{b} \times \bar{\mathbf{p}}') = \bar{\mathbf{p}}^t B \bar{\mathbf{p}}' = 0 \tag{8.6}$$

where B is a skew-symmetric matrix defined by the elements of $\mathbf{b}$: $(b_x,\ b_y,\ b_z)$,

$$B = \begin{pmatrix} 0 & -b_z & b_y \\ b_z & 0 & -b_x \\ -b_y & b_x & 0 \end{pmatrix} \tag{8.7}$$

Using the notations and relations defined so far, we can rewrite the coplanarity equation as

$$\mathbf{p}^t\, \Omega^t\, R^t\, B\, R'\, \Omega'\, \mathbf{p}' = 0 \tag{8.8}$$

We shall call this form the *explicit coplanarity equation*, because the role of each geometric parameters is explicit shown. This becomes more clear if we rewrite this equation in analytical form

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}^t \begin{pmatrix} 1 & 0 & -x_c \\ 0 & 1 & -y_c \\ 0 & 0 & -f \end{pmatrix}^t \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}^t \begin{pmatrix} 0 & -b_z & b_y \\ b_z & 0 & -b_x \\ -b_y & b_x & 0 \end{pmatrix} \begin{pmatrix} r'_{11} & r'_{12} & r'_{13} \\ r'_{21} & r'_{22} & r'_{23} \\ r'_{31} & r'_{32} & r'_{33} \end{pmatrix} \begin{pmatrix} 1 & 0 & -x'_c \\ 0 & 1 & -y'_c \\ 0 & 0 & -f' \end{pmatrix} \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = 0 \tag{8.9}$$

8.2.2 Implicit Coplanarity Equation

Let

$$A = R^t\, B\, R' \tag{8.10}$$

equation (8.8) then becomes

$$\tilde{\mathbf{p}}^t\, A\, \tilde{\mathbf{p}}' = 0 \tag{8.11}$$

or, in expanded form,

$$\begin{pmatrix} x - x_c & y - y_c & -f \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x' - x'_c \\ y' - y'_c \\ -f' \end{pmatrix} = 0 \tag{8.12}$$

where a_{ij}, i, j = 1, 2, 3, are the elements of A. We term equation (8.12) the *first form of the implicit coplanarity equation*, and A the *special coplanarity matrix.*
It is useful at this stage to consider folding $(x_c, y_c, f, x'_c, y'_c, f')$ into A, thereby obtaining an equation which is made up explicit only of the purely measured coordinates of image points. Let

$$D = \Omega^t\, A\, \Omega' \tag{8.13}$$

equation (8.11) then can be rewritten as

$$\mathbf{p}^t\, D\, \mathbf{p}' = 0 \tag{8.14}$$

or, in expanded form,

$$\begin{pmatrix} x & y & 1 \end{pmatrix} \begin{pmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{pmatrix} \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = 0 \tag{8.15}$$

where d_{ij}, i, j = 1, 2, 3, are the elements of D. We term equation (8.15) the *second form of the implicit coplanarity equation*, and D the *general coplanarity matrix.*

The equations (8.12) and (8.15) may be considered as a generalized reformulation of the *relative linear transform* first introduced by Thompson (1959) and then followed by Stefanovic (1973). In the work of Longuet-Higgins (1981), the intrinsic parameters (x_c, y_c, f) and (x'_c, y'_c, f') are assumed known, and a coordinate system equivalent to the second reference system described in section 8.2.3 is adopted. Longuet-Higgins' *essential matrix* is equivalent to the special coplanarity matrix, but in a less general setting. Faugeras et al (1992) considered the more general case in which the intrinsic parameters are unknown, but are fixed across all images (with camera adjustment disabled, as would apply with a frozen mobile camera). A *fundamental matrix* was defined which is essentially equivalent to the above general coplanarity matrix, D, in the case where $(x_c, y_c, f) = (x'_c, y'_c, f')$. Faugeras et al. actually considered a more general camera model with a greater range of linear distortions and associated free parameters than that considered here, but they assume these parameters have already been found by calibrating the camera: they do not attempt to solve for them along with rotation and baseline parameters for the case of two images. The reason we term matrix A in (8.12) and D in (8.15) the *special* and *general coplanarity matrix* respectively is to keep the consistency with the term *coplanarity equation* which is in existence for decades.

8.2.3 Choices of the Reference System

We now explain why a global coordinate system $O-XYZ$ is introduced at the beginning of this section. The reason is simple: because there are different ways of choosing the reference system for the general relative orientation, each with its trade-off.

Reference System 1:

The first choice is to align the global coordinate system with baseline vector and the left (or right) image, such that $O = C$, $O - X = CC'$, and the $O - XZ$ plane contains the left (or right) principal point c. The two-way ambiguity of the $O-Y$ axis's direction may be resolved by adoption of the right-hand rule. We now have that

$$\mathbf{b} = (b, 0, 0)^t, \qquad \alpha = 0 \tag{8.16}$$

where b is the magnitude of $\mathbf{b}$. Because b cannot be determined, the degrees of freedom are now given by $(\beta, \gamma, \alpha', \beta', \gamma', x_c, y_c, f, x'_c, y'_c, f')$.

With this reference system, the explicit coplanarity equation (8.8) becomes

$$\mathbf{p}^t\, \Omega^t\, R^t\, R'\, \Omega'\, \mathbf{p}' = 0 \tag{8.17}$$

where $R = R_\beta\, R_\gamma$. The special coplanarity matrix A of (8.10) becomes

$$A \;=\; R^t\, R' \tag{8.18}$$

This reference system is inspired by the biological vision system. It is neutral and natural to take the baseline as the X-axis. With this reference system, the disparity for a pair of homologous points can be decomposed to horizontal and vertical components. Horizontal disparity can be defined to the orientational difference of the left and right viewing rays in the baseline direction, while the vertical disparity refers to the difference in the direction perpendicular to the plane defined by the baseline and the left principal axis.

Reference System 2:

The second choice is the coincidence of the global and image coordinate systems. Without loss of generality, we choose the left image system. When $O-XYZ$ coincides with $C-\tilde{x}\tilde{y}\tilde{z}$, the left view is defined by the values of the intrinsic parameters (x_c, y_c, f) as well as the equalities

$$R = I, \qquad (X_C, Y_C, Z_C)^T = (0, 0, 0)^T \tag{8.19}$$

where I is the identity matrix. The nature of the right view is fixed by the specification of the intrinsic parameters (x'_c, y'_c, f'), the matrix R' (incorporating the rotation angles α', β', γ'), and the direction of the baseline vector given by

$$\mathbf{B} \;=\; (X_{C'}, Y_{C'}, Z_{C'})^T \tag{8.20}$$

With this reference system, the parameter set for characterising the stereo geometry becomes $(\alpha', \beta', \gamma', b_x, b_y, b_z, x_c, y_c, f, x'_c, y'_c, f')$, where only two of the three components b_x, b_y, b_z need to be determined. Apparently, this reference system is more biased to the left (or right) image.
Note that, for the remainder of this work, we shall assume this second reference system being used. Since we shall no longer deal with both left and right rotation matrices, we henceforth drop the prime and refer to R' as $R = (r_{ij})$, $i, j = 1, 2, 3$, this being the rotation matrix mapping the right image system into the left image system. The special coplanarity matrix A (8.10) thus becomes

$$A = BR \tag{8.21}$$

which is identical to Longuet-Higgins' *essential matrix*.

8.2.4 Properties of Coplanarity Matrices

Similar to the properties of the *essential matrix* and the *fundamental matrix* found by Huang and Faugeras (1989) and Faugeras et al (1992), some obvious properties of the special and general coplanarity matrices A and D can be easily shown,

1. Because the determinant of matrix B is zero, from (8.10) and (8.13), we know

$$|A| = |D| = 0 \tag{8.22}$$

2. From the implicit coplanarity equations (8.11) and (8.14), we know both A and D only determined to a scale factor.
3. From the above two properties, we know the degree of freedom (DoF) of D is 7, i.e.

$$DoF(D) = 7 \tag{8.23}$$

   We also know that A is solely defined by 5 relative orientation parameters, whatever these parameters are chosen, so

$$DoF(A) = 5 \tag{8.24}$$

4. Because the first property (8.22) is a third-order polynomial equation, we know the degree of linear freedom (number of linearly noncorrelated coefficients) (DoLF) is 8, i.e.

$$DoLF(D) = 8 \tag{8.25}$$

Similarity

$$DoLF(A) = 8 \tag{8.26}$$

The third property (8.23) shows that at most only 7 parameters can be solved from the relative geometry of two images using only image measurements. It is another confirmation of seven degrees of freedom in relative stereo geometry known previously [Kruppa, 1913]. That is why we choose seven unknowns (two principal distances and five relative orientation parameters) to define the problem of general relative orientation as the central topic of this work.

It is the fourth property (8.25) and (8.26) that makes it possible to solve for D or A from pure image measurements via a simultaneous collection of implicit coplanarity equations. In addition to these simple properties, we will show more complicated properties by fully exploiting the orthonomality of the rotation matrix R. Using equation (8.21), we obtain a direct relation between A and B,

$$A\,A^t = (B\,R)\,(B\,R)^t = B\,R\,R^t\,B^t = B\,B^t \tag{8.27}$$

Expanding this equation, we have

$$\begin{aligned}
a_{11}^2 + a_{12}^2 + a_{13}^2 &= b_y^2 + b_z^2 && (8.28)\\
a_{21}^2 + a_{22}^2 + a_{23}^2 &= b_x^2 + b_z^2 && (8.29)\\
a_{31}^2 + a_{32}^2 + a_{33}^2 &= b_x^2 + b_y^2 && (8.30)\\
a_{11}a_{21} + a_{12}a_{22} + a_{13}a_{23} &= -b_x b_y && (8.31)\\
a_{11}a_{31} + a_{12}a_{32} + a_{13}a_{33} &= -b_x b_z && (8.32)\\
a_{21}a_{31} + a_{22}a_{32} + a_{23}a_{33} &= -b_y b_z && (8.33)
\end{aligned}$$

Equations (8.28) – (8.30) are simple linear equations of b_x^2, b_y^2, and b_z^2, so

$$\begin{pmatrix} b_x^2 \\ b_y^2 \\ b_z^2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} a_{11}^2 + a_{12}^2 + a_{13}^2 \\ a_{21}^2 + a_{22}^2 + a_{23}^2 \\ a_{31}^2 + a_{32}^2 + a_{33}^2 \end{pmatrix} \tag{8.34}$$

From equations (8.31) – (8.33) we can also solve for b_x^2, b_y^2, and b_z^2 as

$$(b_x^2, b_y^2, b_z^2)^T = \left(\frac{b_{xy}b_{xz}}{b_{yz}}, \frac{b_{xy}b_{yz}}{b_{xz}}, \frac{b_{xz}b_{yz}}{b_{xy}}\right)^T \tag{8.35}$$

where

$$\begin{aligned}
b_{xy} &= -(a_{11}a_{21} + a_{12}a_{22} + a_{13}a_{23}) && (8.36)\\
b_{xz} &= -(a_{11}a_{31} + a_{12}a_{32} + a_{13}a_{33}) && (8.37)\\
b_{yz} &= -(a_{21}a_{31} + a_{22}a_{32} + a_{23}a_{33}) && (8.38)
\end{aligned}$$

Substituting (8.35) into (8.34) gives 3 equations defined purely in terms of the elements of A,

$$(-a_{13}^2 + a_{23}^2 + a_{31}^2 + a_{32}^2 + a_{33}^2 - a_{11}^2 - a_{12}^2 + a_{21}^2 + a_{22}^2)(a_{21}a_{31} + a_{22}a_{32} + a_{23}a_{33}) \\ +2(a_{11}a_{21} + a_{12}a_{22} + a_{13}a_{23})(a_{11}a_{31} + a_{12}a_{32} + a_{13}a_{33}) = 0 \quad (8.39)$$

$$(a_{13}^2 - a_{23}^2 + a_{31}^2 + a_{32}^2 + a_{33}^2 + a_{11}^2 + a_{12}^2 - a_{21}^2 - a_{22}^2)(a_{11}a_{31} + a_{12}a_{32} + a_{13}a_{33}) \\ +2(a_{11}a_{21} + a_{12}a_{22} + a_{13}a_{23})(a_{21}a_{31} + a_{22}a_{32} + a_{23}a_{33}) = 0 \quad (8.40)$$

$$(a_{13}^2 + a_{23}^2 - a_{31}^2 - a_{32}^2 - a_{33}^2 + d_{11}^2 + d_{12}^2 + d_{21}^2 + d_{22}^2)(a_{11}a_{21} + a_{12}a_{22} + a_{13}a_{23}) \\ +2(a_{11}a_{31} + a_{12}a_{32} + a_{13}a_{33})(a_{21}a_{31} + a_{22}a_{32} + a_{23}a_{33}) = 0 \quad (8.41)$$

These will later prove valuable in deriving a direct closed-form solution for solving for two principal distances from the general coplanarity matrix D. Equations (8.39) - (8.41) are not known in the literature, which can only be derived with such elementary manipulations. Due to equation (8.24) and considering the properties in (8.22) and (8.26), we know that one of the three equations (8.39)-(8.41) is redundant. However, this redundancy is useful for robustness.

8.2.5 Outline of the Whole Procedure

In the following sections, we shall show how to recover the 2 principal distances and 5 relative orientation parameters through the implicit and explicit coplanarity equations from pure image measurements. The whole procedure consists of these steps: (1) Solving for the general coplanarity matrix D from eight or more pairs of homologous image points. (2) Solving for the 2 principal distances f and f' from the general coplanarity matrix D via a novel direct closed-form solution. (3) Solving for the special coplanarity matrix A from D and f and f' also via a direct relation. (4) Solving for the baseline vector and rotation matrix from A via a novel direct solution.

8.3 Solving for The Coplanarity Matrix from Image Measurements

The problem of the first step is equivalent to solving the relative linear transform of Thompson (1959), and the essential matrix of Longuet-Higgins (1981), as well as the fundamental matrix of Faugeras (1992). An obvious approach [Thompson, 1959; Stefanovic, 1973; Longuet, 1981; Khlebnikova, 1983] is to transform the implicit coplanarity equation (equivalent) to a normal linear equation by setting one of the implicit coefficients d_{ij}'s to unit. The remaining eight unknowns can thus be solved from the linear equations, using eight or more pairs of homologous image points [Hartley, 1995]. This approach is known to be inefficient and sensitive to noise in image measurements and to the near-singular cases [Brandstätter, 1996]. An alternative approach [Faugeras et al, 1992; Weng et al, 1993] makes use of normalization constraint

$$\sum_{i=1}^{3}\sum_{j=1}^{3} d_{ij} = 1 \qquad (8.42)$$

and solves for the 9 unknowns of D (equivalent) via the singular value decomposition. This second approach is an improvement to the first approach as we no longer need to

arbitrarily set a non-zero d_{ij} to unit, which may actually be close to zero. Again, this approach is known to be still sensitive to image noise and distribution of image points. Another approach [Pan et al, 1995b] is to use an iterative least squares criterion that takes into account the image noise, distribution of image points and constraints. Importantly, the problem of enforcing the constraint (8.22) has been done in INRIA [Zhang et al, 1995]. In addition, the problem of eliminating outliers has been done in Oxford [Torr, 1995].
In the following paragraphs, we present two iterative nonlinear least-square approaches. Many experiments in which moderate image noise is involved show that these approaches give significant improvement to estimating D and are very robust in comparison with those that are solely based on relative linear transform and the singular value decomposition. The rationale of our approach is based on three points:

- the exact use of least-squares criterion,
- iterative numerical approximation, and
- the use of sufficient constraints.

Our approaches also take into account the distribution of homologous image point pairs in each image, and consequently, different weighting factors are properly computed and assigned to each pair of points.
From the process of getting an estimation of the coplanarity matrix D to the final step of recovering all the necessary imaging parameters, improvement has been achieved for each stage. Firstly, the initial estimation of D obtained from the singular value decomposition is iteratively fine tuned using a linear least-squares criterion on the image measurements. This fine tuning process will be given in Section 8.3.1. A further improvement for estimating D is found if the zero determinant of the matrix concerned is also included, this will be discussed in Section 8.3.2.

8.3.1 Iterative Singular Value Decomposition

Given n pairs of homologous image points, a coplanarity equation of the form given in (8.14) is associated with each pair of points. Let (x_i, y_i) and (x'_i, y'_i) be the i-th image point pair then

$$\mathbf{p}_i^T\, D\, \mathbf{p}'_i = M_i\, \mathbf{d} = 0, \tag{8.43}$$

where

$$\begin{aligned} M_i &= (x_i x'_i \quad x_i y'_i \quad x_i \quad y_i x'_i \quad y_i y'_i \quad y_i \quad x'_i \quad y'_i \quad 1) & (8.44)\\ \mathbf{d} &= (d_{11} \quad d_{12} \quad d_{13} \quad d_{21} \quad d_{22} \quad d_{23} \quad d_{31} \quad d_{32} \quad d_{33})^T. & (8.45)\end{aligned}$$

With $n \geq 8$, (8.43) can be written in matrix form as:

$$M\, \mathbf{d} = 0 \tag{8.46}$$

where

$$M = \begin{pmatrix} M_1 \\ M_2 \\ \vdots \\ M_n \end{pmatrix} = \begin{pmatrix} x_1 x'_1 & x_1 y'_1 & x_1 & y_1 x'_1 & y_1 y'_1 & y_1 & x'_1 & y'_1 & 1 \\ x_2 x'_2 & x_2 y'_2 & x_2 & y_2 x'_2 & y_2 y'_2 & y_2 & x'_2 & y'_2 & 1 \\ & & & & \vdots & & & & \\ x_n x'_n & x_n y'_n & x_n & y_n x'_n & y_n y'_n & y_n & x'_n & y'_n & 1 \end{pmatrix}. \tag{8.47}$$

In order to avoid the trivial case where $\mathbf{d} = \mathbf{0}$ is a solution to (8.46) and to also fix the scale factor of $\mathbf{d}$, we also use the normalization constraint (8.42).
Note that this normalization does not uniquely determine the sign of $\mathbf{d}$, so both $\mathbf{d}$ and $-\mathbf{d}$ are valid solutions. For the unique normalization of D, we may set the first non-zero element of $\mathbf{d}$ to be $+1$.
For each pair of homologous image points, an observation equation is available:

$$F_i(\mathbf{d}) = M_i\, \mathbf{d} = 0. \tag{8.48}$$

In general, noise is assumed to exist in each image measurements (x_i, y_i) and (x_i, y_i), so the observation equation can be linearized to

$$F_{x,i}v_{x_i} + F_{y,i}v_{y_i} + F_{x',i}v_{x'_i} + F_{y',i}v_{y'_i} + M_i\mathbf{d} = 0, \tag{8.49}$$

where v_{x_i}, v_{y_i}, $v_{x'_i}$, and $v_{y'_i}$ are, respectively, corrections to x_i, y_i, x'_i, and y'_i; $F_{x,i}$ is the partial derivative of F with respect to x computed for the i-th image point pair:

$$\begin{aligned} F_{x,i} &= x'_i d_{11} + y'_i d_{12} + d_{13} && (8.50)\\ F_{y,i} &= x'_i d_{21} + y'_i d_{22} + d_{23} && (8.51)\\ F_{x',i} &= x_i d_{11} + y_i d_{21} + d_{31} && (8.52)\\ F_{y',i} &= x_i d_{12} + y_i d_{22} + d_{32}. && (8.53) \end{aligned}$$

Let $\mathbf{v}$ be the vector of corrections to the image point observations viz,

$$\mathbf{v} = (v_{x_1} \quad v_{y_1} \quad v_{x'_1} \quad v_{y'_1} \quad \ldots v_{x_n} \quad v_{y_n} \quad v_{x'_n} \quad v_{y'_n})^T, \tag{8.54}$$

then the linearized observation equations can be written as

$$G\mathbf{v} + M\,\mathbf{d} = 0 \tag{8.55}$$

where

$$G = -\begin{pmatrix} F_{x,1} & F_{y,1} & F_{x',1} & F_{y',1} & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & F_{x,2} & F_{y,2} & F_{x',2} & F_{y',2} & \ldots & 0 & 0 & 0 & 0 \\ & & & & & & \vdots & & & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & F_{x,n} & F_{y,n} & F_{x',n} & F_{y',n} \end{pmatrix}. \tag{8.56}$$

The least-squares criterion requires minimizing the objective function

$$\Phi = \mathbf{v}^T W \mathbf{v} - 2\eta^T (G\mathbf{v} + M\mathbf{d}) - \lambda(\mathbf{d}^T\mathbf{d} - 1) \tag{8.57}$$

where W is the $4n \times 4n$ weight matrix of the $4n$ observations x_i, y_i, x'_i, y'_i. Both η and λ are Lagrangian multipliers [Mikhail and Ackermann, 1976]: λ is a scalar coefficient, and η is an n-vector (recalling that n is the number of homologous image point pairs).
In general, image observations are assumed to be uncorrelated, so W is diagonal:

$$W = \begin{pmatrix} w_{x_1} & 0 & \ldots & 0 \\ 0 & w_{y_1} & \ldots & 0 \\ & \vdots & \ddots & \\ 0 & 0 & \ldots & w_{y'_n} \end{pmatrix}. \tag{8.58}$$

Since the first derivative of Φ vanishes when Φ attains its minimum, taking the partial derivatives of Φ in (8.57) and equating them to zero gives

$$\frac{\partial \Phi}{\partial \mathbf{v}} = 2\mathbf{v}^T W - 2\eta^T G = 0 \tag{8.59}$$

$$\frac{\partial \Phi}{\partial \mathbf{d}} = -2\eta^T M - 2\lambda \mathbf{d}^T = 0. \tag{8.60}$$

need to ensure the second derivatives are positive?
From (8.59), we have

$$\mathbf{v} = W^{-1} G^T \eta. \tag{8.61}$$

Applying this to (8.55) yields

$$\eta = -(GW^{-1}G^T)^{-1} M \mathbf{d}, \tag{8.62}$$

and substituting the above to (8.60) gives

$$(M^T (GW^{-1}G^T)^{-1} M)\, \mathbf{d} = \lambda \mathbf{d}. \tag{8.63}$$

From this equation we see that λ is an eigenvalue of the data matrix $M^T(GW^{-1}G^T)^{-1}M$, and $\mathbf{d}$ is an eigenvector corresponding to this eigenvalue.
Under the least-squares criterion, we seek

$$\begin{aligned} \min \mathbf{v}^T W \mathbf{v} &= \min (W^{-1}G^T\eta)^T W (W^{-1}G^T\eta) \\ &= \min \eta^T (GW^{-1}G^T)\eta \\ &= \min \mathbf{d}^T M^T (GW^{-1}G^T)^{-1}(GW^{-1}G^T)(GW^{-1}G^T)^{-1} M\mathbf{d} \\ &= \min \mathbf{d}^T M^T (GW^{-1}G^T)^{-1} M\mathbf{d} \\ &= \min \lambda \mathbf{d}^T \mathbf{d} \\ &= \min \lambda. \end{aligned} \tag{8.64}$$

The problem then becomes that of finding the smallest eigenvalue λ of the data matrix $M^T(GW^{-1}G^T)^{-1}M$.
Note that if W is diagonal as shown in (8.58), $GW^{-1}G^T$ is also diagonal and its i-th diagonal element is computed as

$$\frac{1}{w_{x_i}} F_{x,i}^2 + \frac{1}{w_{y_i}} F_{y,i}^2 + \frac{1}{w_{x'_i}} F_{x',i}^2 + \frac{1}{w_{y'_i}} F_{y',i}^2. \tag{8.65}$$

The solution to (8.63) requires an iterative nonlinear procedure because the computation of G involves an approximate value of $\mathbf{d}$. To begin the procedure, we first solve for $\mathbf{d}$ from

$$(M^T M)\, \mathbf{d} = \lambda \mathbf{d}. \tag{8.66}$$

This is equivalent to a least-squares solution where the observation equation (8.49) is simplified to

$$v_i + M_i \mathbf{d} = 0 \tag{8.67}$$

where v_i is the corrections to the fictitious observation $M_i\mathbf{d}$ whose ideal value should be zero.

This simplified one-step least-squares solution in (8.66) has been used by Faugeras *et al.* (1992) and Weng *et al.* (1993). The iterative least-squares solution proposed in (8.63) is an improvement, as it takes into consideration the distribution structure of the homologous image points and the uncertainty involved in the image point measurements. Numerical tests show that the first iteration of (8.63) improves the estimation of D at the first significant digit. For subsequent iterations the improvement diminishes.

In fact, a proper normalization to the image measurements is found to reduce the numerical sensitivity associated with estimating the coplanarity matrix D. Let k and k' be two known normalization factors applied to the left and right image points, (8.15) then becomes

$$\begin{pmatrix} x/k & y/k & 1 \end{pmatrix} \begin{pmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{pmatrix} \begin{pmatrix} x'/k' \\ y'/k' \\ 1 \end{pmatrix} = 0. \tag{8.68}$$

This method of normalization for improving the estimation of D has also been reported by Hartley (1995). While Hartley scaled the image coordinates such that the average distance of image points to the centroid of the image point cloud in each image is $\sqrt{2}$, our normalization factors k and k' are determined so that the (average) first 8 elements of M_i in (8.44) are distributed around 1. Observation on (8.44) reveals that the last element of M_i is a fixed constant value and thus our choice of the normalization factors.

Note that the normalization factors k and k' applied to the left and right image coordinates also scale the principal distances f and f' accordingly. The final estimation of f and f' must therefore be inversely scaled.

In summary, the iterative singular value decomposition algorithm given in this section can be briefly described as follows:

1. get an estimate $\mathbf{d}_0$ of $\mathbf{d}$ from (8.66).
2. set $t = 1$.
3. compute the diagonal matrix $(GW^{-1}G^T)^{-1}$. The elements of the diagonal matrix $(GW^{-1}G^T)$ are defined in (8.65).
4. obtain an estimate $\mathbf{d}_t$ of $\mathbf{d}$ from (8.63).
5. terminate the procedure if $||\mathbf{d}_t - \mathbf{d}_{t-1}||$ is less than a prespecified threshold.
6. otherwise, increment t by 1 and go back to step 3.

8.3.2 Iterative Linearized Least Squares

In this section we consider incorporating the zero determinant constraint of D to the least squares method. Two constraints arise from the two properties of D listed in section 8.2.4. In principle, both constraints should be used when estimating D, but most previous research in this area, such as the simple linear method and the singular value decomposition method, had not taken into account this second property. The resultant effect is that, due to image noise, the smallest singular value of D may not vanish and so D is non-singular. It was only very recent that this zero determinant constraint of D is explicitly enforced when this matrix is estimated. For instance, Faugeras et al. (1992) used the following representation for the coplanarity matrix:

$$D = \begin{pmatrix} d_1 & d_2 & d_3 \\ d_4 & d_5 & d_6 \\ d_7 d_1 + d_8 d_4 & d_7 d_2 + d_8 d_5 & d_7 d_3 + d_8 d_6 \end{pmatrix} \tag{8.69}$$

to constrain the third row of D to be spanned by the first two.

This representation also eliminates one unknown from the matrix, which was estimated using unconstrained minimization. To ensure that $|D| = 0$, Hartley (1995), on the other hand, applied the singular value decomposition to the initially estimated D_0, i.e. $D_0 = U\mathrm{diag}(r, s, t)\, V^T$ where both U and V are orthogonal matrices, and $\mathrm{diag}(r, s, t)$, with $r \geq s \geq t$, is a diagonal matrix containing the singular values of D. From this decomposition, D_0 was superseded by a new matrix D_1 for subsequent computation, where D_1 is defined as $U\mathrm{diag}(r, s, 0)\, V^T$.

We have taken an approach different from both Faugeras et al. and Hartley's to enforce this zero determinant constraint into our iterative linear least-squares method. Let us define two functions expressing the two constraints on D:

$$\phi(\mathbf{d}) = \mathbf{d}^T \mathbf{d} - 1 = 0 \tag{8.70}$$

and

$$\psi(\mathbf{d}) = |D| = 0. \tag{8.71}$$

With these two constraints, the objective function of the least-squares criterion can be defined as

$$\Phi = \mathbf{v}^T W \mathbf{v} - 2\eta^T (G\mathbf{v} + M\mathbf{d}) - \lambda(\mathbf{d}^T \mathbf{d} - 1) - \omega |D|, \tag{8.72}$$

where η, λ, and ω are all Lagrangian multipliers: η is an n-vector; both λ and ω are scalars.

As the function ψ is a cubic equation of the elements of $\mathbf{d}$, direct partial derivatives of Φ with respect to $\mathbf{v}$ and $\mathbf{d}$ will not lead to a closed-form solution, such as the singular value decomposition of (8.63). In fact, this will lead to 9 quadratic equations in the elements of $\mathbf{d}$, for which there is no general solution in mathematics.

A natural way to overcome this difficulty is to linearize both constraints in terms of the correction vector $\Delta\mathbf{d}$ to the elements of $\mathbf{d}$. The two constraints can be linearized to

$$\phi(\mathbf{d}, \Delta\mathbf{d}) = 2\mathbf{d}\Delta\mathbf{d} + (\mathbf{d}^T\mathbf{d} - 1) = 0 \tag{8.73}$$

$$\psi(\mathbf{d}, \Delta\mathbf{d}) = \left(\frac{\partial\psi}{\partial\mathbf{d}}\right)^T \Delta\mathbf{d} + |D| = 0, \tag{8.74}$$

which can be put into matrix form as follows:

$$N\Delta\mathbf{d} + U = 0 \tag{8.75}$$

where

$$N = \begin{pmatrix} 2d_{11} & 2d_{12} & 2d_{13} & 2d_{21} & 2d_{22} & 2d_{23} & 2d_{31} & 2d_{32} & 2d_{33} \\ \frac{\partial\psi}{\partial d_{11}} & \frac{\partial\psi}{\partial d_{12}} & \frac{\partial\psi}{\partial d_{13}} & \frac{\partial\psi}{\partial d_{21}} & \frac{\partial\psi}{\partial d_{22}} & \frac{\partial\psi}{\partial d_{23}} & \frac{\partial\psi}{\partial d_{31}} & \frac{\partial\psi}{\partial d_{32}} & \frac{\partial\psi}{\partial d_{33}} \end{pmatrix} \tag{8.76}$$

$$U = \begin{pmatrix} \mathbf{d}^T\mathbf{d} - 1 \\ |D| \end{pmatrix}. \tag{8.77}$$

The original observation equation (8.55) now becomes

$$G\,\mathbf{v} + M\,\Delta\mathbf{d} + M\,\mathbf{d} = 0. \tag{8.78}$$

With the above linearization, the objective function Φ can be redefined as

$$\Phi = \mathbf{v}^T W \mathbf{v} - 2\eta^T (G\mathbf{v} + M\Delta\mathbf{d} + M\mathbf{d}) - 2\kappa^T (N\Delta\mathbf{d} + U), \tag{8.79}$$

where η and κ are two Lagrangian multipliers, with η being an n-vector and κ a 2-vector. The partial derivative of Φ with respect to $\mathbf{v}$ is exactly the same as those given in (8.59). Applying (8.61) into (8.78), we have

$$(GW^{-1}G^T)\eta + M\Delta\mathbf{d} + M\mathbf{d} = 0. \tag{8.80}$$

So,

$$\eta = -(GW^{-1}G^T)^{-1}(M\Delta\mathbf{d} + M\mathbf{d}). \tag{8.81}$$

Taking the partial derivative of Φ with respect to $\Delta\mathbf{d}$, we have

$$\frac{\partial\Phi}{\partial\Delta\mathbf{d}} = -2\eta^T M - 2\kappa^T N = 0. \tag{8.82}$$

This can be rewritten as

$$M^T\eta + N^T\kappa = 0. \tag{8.83}$$

Substituting (8.81) into (8.83), we have

$$-\,M^T(GW^{-1}G^T)^{-1}(M\Delta\mathbf{d} + M\mathbf{d}) + N^T\kappa = 0. \tag{8.84}$$

Let

$$Q = M^T(GW^{-1}G^T)^{-1}M. \tag{8.85}$$

From (8.84) we derive

$$\Delta\mathbf{d} = Q^{-1}N^T\kappa - \mathbf{d}. \tag{8.86}$$

Substituting (8.86) into (8.75), we have

$$NQ^{-1}N^T\kappa - N\mathbf{d} + U = 0. \tag{8.87}$$

Let

$$S = NQ^{-1}N^T. \tag{8.88}$$

From (8.87), we derive

$$\kappa = S^{-1}(N\mathbf{d} - U). \tag{8.89}$$

After κ is solved from (8.89), the unknown vector $\Delta\mathbf{d}$ can be determined via (8.86). This iterative procedure can be summarised as follows:

1. get an estimate $\mathbf{d}_0$ of $\mathbf{d}$ from (8.66).

2. set $t = 1$.

3. compute:
   matrix N as defined in (8.76),
   matrix Q as defined in (8.85),
   matrix S as defined in (8.88),
   vector κ as defined in (8.89), and
   vector $\Delta\mathbf{d}$ as defined in (8.86).

4. set $\mathbf{d}_t = \mathbf{d}_{t-1} + \Delta\mathbf{d}$.

5. terminate the iteration procedure if $||\mathbf{d}_t - \mathbf{d}_{t-1}||$ is less than a prespecified threshold.

6. otherwise, increment t by 1 and go back to step 3.

8.4 Solving for the Two Principal Distances

8.4.1 Basic Relation and Constraint

In practical applications, we want to include two principal distances f and f' as unknowns, given only seven degrees of freedom, we must assume the principal point coordinates x_c and y_c to be known *a priori* (e.g. detected by using fiducial marks for a metric camera), or to be sufficiently close to zero. In order to solve for two principal distances f and f' from the solved general coplanarity matrix D, we need to first defactorize the principal points (x_c, y_c) and (x'_c, y'_c) out from D. Using the equations (2.8), and (8.13), we obtain

$$D = \Omega^t A \, \Omega' = \Omega_c^t \, \Omega_f^t \, A \, \Omega'_f \, \Omega'_c \tag{8.90}$$

where

$$\Omega_c = \begin{pmatrix} 1 & 0 & -x_c \\ 0 & 1 & -y_c \\ 0 & 0 & 1 \end{pmatrix}, \qquad \Omega_f = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -f \end{pmatrix} \tag{8.91}$$

Let

$$\bar{D} = \Omega_f^t \, A \, \Omega'_f \tag{8.92}$$

$\bar{D}$ is a special version of D,

$$\bar{D} = D, \text{ if } (x_c, y_c) = (x'_c, y'_c) = (0, 0) \tag{8.93}$$

In general when $(x_c, y_c), (x'_c, y'_c)$ are nonzero, but known *a priori*, $\bar{D}$ can be solved from D via (8.90) as

$$\bar{D} = \Omega_c^{-t} \, D \Omega_c'^{-1} \tag{8.94}$$

Note that $\bar{D}$ involves purely the 7 degrees of freedom in the coplanarity constraint, which correspond to 7 unknowns (2 principal distances and 5 relative orientation parameters)

in the general relative orientation, we may simply call $\bar{D}$ the *coplanarity matrix*. $\bar{D}$ has the same properties of D as listed in the subsection 8.2.4.
In fact, we can also capture x_c and y_c into the image measurements. The implicit coplanarity equation (8.15) can be written as

$$\begin{pmatrix} x - x_c & y - y_c & 1 \end{pmatrix} \bar{D} \begin{pmatrix} x' - x'_c \\ y' - y'_c \\ 1 \end{pmatrix} = 0 \tag{8.95}$$

$\bar{D}$ defined by (8.95) can still be solved using the general approaches for solving D, but we need to use $x - x_c$ and $y - y_c$ instead of x and y. In the remainder of this work, we shall assume $\bar{D}$ is solved either via (8.94) or directly via (8.95).
The relation between $\bar{D} = (\bar{d}_{ij})$ and $A = (a_{ij})$ is

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} \bar{d}_{11} & \bar{d}_{12} & \frac{\bar{d}_{13}}{f'} \\ \bar{d}_{21} & \bar{d}_{22} & \frac{\bar{d}_{23}}{f'} \\ \frac{\bar{d}_{31}}{f} & \frac{\bar{d}_{32}}{f} & \frac{\bar{d}_{33}}{f\,f'} \end{pmatrix} \tag{8.96}$$

8.4.2 Algebraic Equations of Two Principal Distances

Applying expression (8.96) into (8.39) – (8.41), 3 equations in 2 unknowns f and f' are obtained

$$h_{i1} + h_{i2} f^2 + h_{i3} f'^2 + h_{i4} f'^4 + h_{i5} f^2 f'^2 + h_{i6} f^2 f'^4 = 0, \qquad (i = 1, 2, 3) \tag{8.97}$$

where the h_{ij}'s are coefficients directly computable from the elements $\bar{d}_{ij}$'s, given by

$$\begin{aligned}
h_{11} &= \bar{d}_{23}\bar{d}_{33}^3 && (8.98)\\
h_{12} &= (\bar{d}_{13}^2 + +\bar{d}_{23}^2)\bar{d}_{23}\bar{d}_{33} && (8.99)\\
h_{13} &= (\bar{d}_{31}^2 + \bar{d}_{32}^2)\bar{d}_{23}\bar{d}_{33} + (\bar{d}_{21}\bar{d}_{31} + \bar{d}_{22}\bar{d}_{32})\bar{d}_{33}^2 && (8.100)\\
h_{14} &= (\bar{d}_{31}^2 + \bar{d}_{32}^2)(\bar{d}_{21}\bar{d}_{31} + \bar{d}_{22}\bar{d}_{32}) && (8.101)\\
h_{15} &= (\bar{d}_{22}^2 - \bar{d}_{11}^2 - \bar{d}_{12}^2 + \bar{d}_{21}^2)\bar{d}_{23}\bar{d}_{33} + (-\bar{d}_{13}^2 + \bar{d}_{23}^2)(\bar{d}_{21}\bar{d}_{31} + \bar{d}_{22}\bar{d}_{32}) \\
&\quad +2(\bar{d}_{11}\bar{d}_{21} + \bar{d}_{12}\bar{d}_{22})\bar{d}_{13}\bar{d}_{33} + 2(\bar{d}_{11}\bar{d}_{31} + \bar{d}_{12}\bar{d}_{32})\bar{d}_{13}\bar{d}_{23} && (8.102)\\
h_{16} &= (\bar{d}_{22}^2 - \bar{d}_{11}^2 - \bar{d}_{12}^2 + \bar{d}_{21}^2)(\bar{d}_{21}\bar{d}_{31} + \bar{d}_{22}\bar{d}_{32}) \\
&\quad +2(\bar{d}_{11}\bar{d}_{21} + \bar{d}_{12}\bar{d}_{22})(\bar{d}_{11}\bar{d}_{31} + \bar{d}_{12}\bar{d}_{32}) && (8.103)\\
h_{21} &= \bar{d}_{13}\bar{d}_{33}^3 && (8.104)\\
h_{22} &= (\bar{d}_{13}^2 + \bar{d}_{23}^2)\bar{d}_{13}\bar{d}_{33} && (8.105)\\
h_{23} &= (\bar{d}_{31}^2 + \bar{d}_{32}^2)\bar{d}_{13}\bar{d}_{33} + (\bar{d}_{11}\bar{d}_{31} + \bar{d}_{12}\bar{d}_{32})\bar{d}_{33}^2 && (8.106)\\
h_{24} &= (\bar{d}_{31}^2 + \bar{d}_{32}^2)(\bar{d}_{11}\bar{d}_{31} + \bar{d}_{12}\bar{d}_{32}) && (8.107)\\
h_{25} &= (\bar{d}_{11}^2 + \bar{d}_{12}^2 - \bar{d}_{21}^2 - \bar{d}_{22}^2)\bar{d}_{13}\bar{d}_{33} + (\bar{d}_{13}^2 - \bar{d}_{23}^2)(\bar{d}_{11}\bar{d}_{31} + \bar{d}_{12}\bar{d}_{32}) \\
&\quad +2(\bar{d}_{11}\bar{d}_{21} + \bar{d}_{12}\bar{d}_{22})\bar{d}_{23}\bar{d}_{33} + 2(\bar{d}_{21}\bar{d}_{31} + \bar{d}_{22}\bar{d}_{32})\bar{d}_{13}\bar{d}_{23} && (8.108)\\
h_{26} &= (\bar{d}_{11}^2 + \bar{d}_{12}^2 - \bar{d}_{21}^2 - \bar{d}_{22}^2)(\bar{d}_{11}\bar{d}_{31} + \bar{d}_{12}\bar{d}_{32}) \\
&\quad +2(\bar{d}_{11}\bar{d}_{21} + \bar{d}_{12}\bar{d}_{22})(\bar{d}_{21}\bar{d}_{31} + \bar{d}_{22}\bar{d}_{32}) && (8.109)\\
h_{31} &= \bar{d}_{13}\bar{d}_{23}\bar{d}_{33}^2 && (8.110)
\end{aligned}$$

$$h_{32} = (\bar{d}_{13}^2 + \bar{d}_{23}^2)\bar{d}_{13}\bar{d}_{23} \tag{8.111}$$
$$h_{33} = -(\bar{d}_{31}^2 + \bar{d}_{32}^2)\bar{d}_{13}\bar{d}_{23} - (\bar{d}_{11}\bar{d}_{21} + \bar{d}_{12}\bar{d}_{22})\bar{d}_{33}^2 + 2(\bar{d}_{11}\bar{d}_{31} + \bar{d}_{12}\bar{d}_{32})\bar{d}_{23}\bar{d}_{33} + 2(\bar{d}_{21}\bar{d}_{31} + \bar{d}_{22}\bar{d}_{32})\bar{d}_{13}\bar{d}_{33} \tag{8.112}$$
$$h_{34} = -(\bar{d}_{31}^2 + \bar{d}_{32}^2)(\bar{d}_{11}\bar{d}_{21} + \bar{d}_{12}\bar{d}_{22}) + 2(\bar{d}_{11}\bar{d}_{31} + \bar{d}_{12}\bar{d}_{32})(\bar{d}_{21}\bar{d}_{31} + \bar{d}_{22}\bar{d}_{32}) \tag{8.113}$$
$$h_{35} = (\bar{d}_{11}^2 + \bar{d}_{12}^2 + \bar{d}_{21}^2 + \bar{d}_{22}^2)\bar{d}_{13}\bar{d}_{23} + (\bar{d}_{13}^2 + \bar{d}_{23}^2)(\bar{d}_{11}\bar{d}_{21} + \bar{d}_{12}\bar{d}_{22}) \tag{8.114}$$
$$h_{36} = (\bar{d}_{11}^2 + \bar{d}_{12}^2 + \bar{d}_{21}^2 + \bar{d}_{22}^2)(\bar{d}_{11}\bar{d}_{21} + \bar{d}_{12}\bar{d}_{22}) \tag{8.115}$$

From each of equations (8.97), f^2 can be expressed in terms of f' viz:

$$f^2 = f_i^2 \equiv -\frac{h_{i1} + h_{i3}f'^2 + h_{i4}f'^4}{h_{i2} + h_{i5}f'^2 + h_{i6}f'^4}, \qquad (i = 1,2,3) \tag{8.116}$$

In theory, we have that $f_1^2 = f_2^2$, $f_2^2 = f_3^2$, and $f_1^2 = f_3^2$. Imposing each of these constraints in turn, then, after setting

$$q = f'^2 \tag{8.117}$$

we obtain three cubic algebraic equations in q given by

$$s_{i1} + s_{i2}q + s_{i3}q^2 + s_{i4}q^3 = 0, \qquad (i = 1,2,3) \tag{8.118}$$

where the s_{ij}'s are coefficients directly computable from h_{ij}'s, given by

$$s_{11} = -h_{11}h_{25} + h_{12}h_{23} - h_{13}h_{22} + h_{15}h_{21} \tag{8.119}$$
$$s_{12} = -h_{11}h_{26} + h_{12}h_{24} - h_{13}h_{25} - h_{14}h_{22} + h_{15}h_{23} + h_{16}h_{21} \tag{8.120}$$
$$s_{13} = -h_{13}h_{26} - h_{14}h_{25} + h_{15}h_{24} + h_{16}h_{23} \tag{8.121}$$
$$s_{14} = -h_{14}h_{26} + h_{16}h_{24} \tag{8.122}$$
$$s_{21} = -h_{11}h_{35} + h_{12}h_{33} - h_{13}h_{32} + h_{15}h_{31} \tag{8.123}$$
$$s_{22} = -h_{11}h_{36} + h_{12}h_{34} - h_{13}h_{35} - h_{14}h_{32} + h_{15}h_{33} + h_{16}h_{31} \tag{8.124}$$
$$s_{23} = -h_{13}h_{36} - h_{14}h_{35} + h_{15}h_{34} + h_{16}h_{33} \tag{8.125}$$
$$s_{24} = -h_{14}h_{36} + h_{16}h_{34} \tag{8.126}$$
$$s_{31} = -h_{21}h_{35} + h_{22}h_{33} - h_{23}h_{32} + h_{25}h_{31} \tag{8.127}$$
$$s_{32} = -h_{21}h_{36} + h_{22}h_{34} - h_{23}h_{35} - h_{24}h_{32} + h_{25}h_{33} + h_{26}h_{31} \tag{8.128}$$
$$s_{33} = -h_{23}h_{36} - h_{24}h_{35} + h_{25}h_{34} + h_{26}h_{33} \tag{8.129}$$
$$s_{34} = -h_{24}h_{36} + h_{26}h_{34} \tag{8.130}$$

Equations (8.118) are the central result of this paper. Each of the equations (8.118) is cubic in q, and can therefore be solved in closed-form. Once q is solved, f' and f are then solved from (8.117) and (8.116). And the special coplanarity matrix A is obtained from (8.96).

8.4.3 The Case of Two Equal Principal Distances

In practical applications, there is an usual case where two principal distances are equal, i.e.

$$f = f' \tag{8.131}$$

This occurs in standard aerial photogrammetry and motion vision with a 'frozen' mobile camera. In this case, the algebraic equations of (8.97) becomes

$$h_{i1} + (h_{i2} + h_{i3})f^2 + (h_{i4} + h_{i5})f^4 + h_{i6}f^6 = 0 \qquad (i = 1, 2, 3) \tag{8.132}$$

These are cubic equations in an unknown f^2, so f^2 can be solved in closed-form.

8.4.4 Cases of Degeneracy

As the degree of freedom of $\bar{D}$ is 7, it appears that if one of the elements of $\bar{D}$ constantly equals zero, we can not solve for all the 7 parameters from $\bar{D}$. This becomes clear when we consider the relations between $\bar{D}$ and A, and A and B, R.
Expanding equation (8.21) gives

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} b_y r_{31} - b_z r_{21} & b_y r_{32} - b_z r_{22} & b_y r_{33} - b_z r_{23} \\ b_z r_{11} - b_x r_{31} & b_z r_{12} - b_x r_{32} & b_z r_{13} - b_x r_{33} \\ b_x r_{21} - b_y r_{11} & b_x r_{22} - b_y r_{12} & b_x r_{23} - b_y r_{13} \end{pmatrix} \tag{8.133}$$

By fixing a $\bar{d}_{ij}$ to zero, for a particular index of (i, j), from relation (8.96), it implies

$$\bar{d}_{ij} \equiv 0 \quad \Longrightarrow \quad a_{ij} \equiv 0 \tag{8.134}$$

That means there is a constant dependency between the baseline vector (b_x, b_y, b_z) and the rotation angles (α, β, γ). Obviously, there are 9 such cases of degeneracy. Not every case is meaningful in practical setup of the stereo geometry. However, there is at least one practical case of degeneracy

$$\bar{d}_{33} \equiv 0 \quad \Longrightarrow \quad a_{33} \equiv 0 \quad \Longrightarrow \quad b_x r_{23} - b_y r_{13} \equiv 0 \tag{8.135}$$

This degenerate case corresponds to the coplanarity of two principal axes which is the standard case of biological and robot stereo setup. In standard aerial photogrammetry, two stereo image planes tend to be parallel, which also corresponds to the coplanarity of two principal axes. That two principal axes are coplanar naturally implies they are coplanar with the baseline vector

$$[\mathbf{b} \quad \overline{Cc} \quad \overline{C'c'}] = \begin{vmatrix} b_x & b_y & b_z \\ 0 & 0 & -f \\ -r_{13}f' & -r_{23}f' & -r_{33}f' \end{vmatrix} \equiv 0 \tag{8.136}$$

This can be reduced to

$$\begin{vmatrix} b_x & b_y \\ r_{13} & r_{23} \end{vmatrix} = b_x r_{23} - b_y r_{13} \equiv 0 \tag{8.137}$$

This shows that $\bar{d}_{33} \equiv 0$ corresponds to a particular geometry of stereo setup, which is the standard case of robot and biological vision. Further to show the constraint on the explicit parameters, replacing r_{13} and r_{23} by their trigonometric functions of rotation angles defined by (2.2), equation (8.137) can be expanded as

$$b_x r_{23} - b_y r_{13} = b_x \sin\alpha \cos\beta + b_y \sin\beta \equiv 0 \tag{8.138}$$

This shows that this particular degenerate case corresponds to a constant dependence between two of the baseline components and two of the rotation angles: b_x, b_y, α, β, not involving b_z and γ.
This degenerate case can also be confirmed by using the algebraic equations in f and f' of (8.97). When $\bar{d}_{33} \equiv 0$, the three equations in f and f' of (8.97) reduce to only one equation. Consequently, there is no unique solution.
In the standard case of robot and biological vision, only a vergence angle β is allowed to be variable, so

$$\alpha = \gamma = 0 \tag{8.139}$$

If the two cameras are 'frozen' to have equal principal distances $f = f'$, the unique principal distance f is then solvable. This is a special case of equation (8.132). Numerical tests show that by setting α to a tiny nonzero angle (e.g. $1/\pi$), it is sufficiently robust to avoid this degenerate case, so the two different focal lengths can be solved.

8.5 Solving For The Relative Baseline and Rotation

8.5.1 Two Symmetric Solutions for the Baseline Vector

After the two principal distances f and f', and a_{ij}'s are solved, b_x^2, b_y^2, and b_z^2 can also be determined via equations (8.28)-(8.30) as

$$\begin{pmatrix} b_x^2 \\ b_y^2 \\ b_z^2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} a_{11}^2 + a_{12}^2 + a_{13}^2 \\ a_{21}^2 + a_{22}^2 + a_{23}^2 \\ a_{31}^2 + a_{32}^2 + a_{33}^2 \end{pmatrix} \tag{8.140}$$

The signs of b_x, b_y, and b_z can then be determined by using equations (8.31)-(8.33). Two symmetric solutions $\mathbf{b}_1 = (b_{x1}, b_{y1}, b_{z1})$ and $\mathbf{b}_2 = (b_{x2}, b_{y2}, b_{z2})$ will be obtained, i.e.

$$\mathbf{b}_1 = -\mathbf{b}_2 \tag{8.141}$$

We may resolve the sign ambiguity as follows. From equations (8.31)-(8.33) and (8.36)-(8.38), we know

$$b_{xy} = b_x b_y \tag{8.142}$$

$$b_{xz} = b_x b_z \tag{8.143}$$

$$b_{yz} = b_y b_z \tag{8.144}$$

All combinations of the signs of b_x, b_y, b_z, b_{xy}, b_{xz}, and b_{yz} can be tabulated as

Table 1

b_x	b_y	b_z	b_{xy}	b_{xz}	b_{yz}
+	+	+	+	+	+
+	+	−	+	−	−
+	−	+	−	+	−
+	−	−	−	−	+
−	+	+	−	−	+
−	+	−	−	+	−
−	−	+	+	−	−
−	−	−	+	+	+

From this table, for any combination of the known signs of b_{xy}, b_{xz}, and b_{yz}, we can select a unique symmetric pair of baseline vectors.

At this stage, we can not determine which of the two solutions for the baseline vector is uniquely valid in practice. However, with the two baseline solutions and the given A matrix, we can solve for two different rotation matrices R_1 and R_2, and the corresponding sets of rotation angles $(\alpha_1, \beta_1, \gamma_1)$ and $(\alpha_2, \beta_2, \gamma_2)$. This can be done in the following way.

8.5.2 Two Symmetric Solutions for the Orientation Matrix and Angles

Given the special coplanarity matrix A and a candidate baseline solution, the rotation matrix R can be solved via equation (8.21). As $|B| = 0$, however, R cannot be solved directly from this relation. The orthonormality of R needs to be exploited, and there are two possible approaches.

The first approach is to represent R in terms of three independent elements. To avoid using explicit rotation angles and trigonometric functions, we can express R in an algebraic form. Let S be an skew-symmetric matrix constructed solely from three independent parameters a, b, c:

$$S = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix} \tag{8.145}$$

It can then be proved that any orthonomal matrix R can be constructed by using S as

$$R = (I + S)(I - S)^{-1} \tag{8.146}$$

where I is the 3×3 identity matrix. Expanding (8.146) leads to

$$R = \frac{1}{1 + a^2 + b^2 + c^2} \begin{pmatrix} 1 + a^2 - b^2 - c^2 & -2c + 2ab & 2b + 2ac \\ 2c + 2ab & 1 - a^2 + b^2 - c^2 & -2a + 2bc \\ -2b + 2ac & 2a + 2bc & 1 - a^2 - b^2 + c^2 \end{pmatrix} \tag{8.147}$$

By applying this representation into (8.133), we obtain 9 quadratic equations that are overconstrained for solving the three unknowns a, b, and c. Although this approach has the advantage that a minimum number of unknowns is involved, it leads to highly nonlinear equations which are difficult to solve.

The second approach is to use 6 unknowns, e.g. r_{11}, r_{21}, r_{31}, r_{21}, r_{22}, r_{23}. As R is orthonormal, and $|R| = 1$, any element r_{ij} can be represented by its cofactor, e.g.

$$r_{13} = \begin{vmatrix} r_{21} & r_{22} \\ r_{31} & r_{32} \end{vmatrix}, \quad r_{23} = -\begin{vmatrix} r_{11} & r_{12} \\ r_{31} & r_{32} \end{vmatrix}, \quad r_{33} = \begin{vmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{vmatrix} \tag{8.148}$$

By using this property of R, from (8.133) we obtain

$$\begin{pmatrix} b_x^2 & b_x b_y & b_x b_z \\ b_x b_y & b_y^2 & b_y b_z \\ b_x b_z & b_y b_z & b_z^2 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \\ r_{31} & r_{32} \end{pmatrix} = \begin{pmatrix} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} & -\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} \\ -\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} \\ \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} & -\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} \end{pmatrix} \tag{8.149}$$

Directly from relation (8.133) we also obtain

$$\begin{pmatrix} 0 & -b_z & b_y \\ b_z & 0 & -b_x \\ -b_y & b_x & 0 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \\ r_{31} & r_{32} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \tag{8.150}$$

Although each of the coefficient matrices in (8.149) and (8.150) is singular, particular combinations of their equations can lead to nonsingular coefficient matrices.
If $b_x \neq 0$, we can use

$$\begin{pmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \\ r_{31} & r_{32} \end{pmatrix} = \begin{pmatrix} b_z & 0 & -b_x \\ -b_y & b_x & 0 \\ b_x^2 & b_x b_y & b_x b_z \end{pmatrix}^{-1} \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \\ \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} & -\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} \end{pmatrix} \tag{8.151}$$

The determinant of the coefficient matrix to be inverted is $b_x^2(b_x^2 + b_y^2 + b_z^2)$, so if $b_x \neq 0$, the inverse is guaranteed to exist.
Similarly, if $b_y \neq 0$, we can use

$$\begin{pmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \\ r_{31} & r_{32} \end{pmatrix} = \begin{pmatrix} 0 & -b_z & b_y \\ -b_y & b_x & 0 \\ b_x b_y & b_y^2 & b_y b_z \end{pmatrix}^{-1} \begin{pmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \\ -\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} \end{pmatrix} \tag{8.152}$$

If $b_z \neq 0$, we can use

$$\begin{pmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \\ r_{31} & r_{32} \end{pmatrix} = \begin{pmatrix} 0 & -b_z & b_y \\ b_z & 0 & -b_x \\ b_x b_z & b_y b_z & b_z^2 \end{pmatrix}^{-1} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} & -\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} \end{pmatrix} \tag{8.153}$$

After r_{ij}'s, $i = 1, 2, 3$; $j = 1, 2$, are solved, the remaining three elements r_{13}, r_{23}, and r_{33} can be computed via (8.148). After R is solved, the three explicit rotation angles about the three principal axes can be determined through simple inverse trigonometric functions from R.

8.5.3 Determining The Unique Baseline and Orientation

Given the special coplanarity matrix A, two symmetric solutions of the baseline vector $\mathbf{b}_1, \mathbf{b}_2$ and the corresponding rotation matrices $R_1(\alpha_1, \beta_1, \gamma_1)$ and $R_2(\alpha_2, \beta_2, \gamma_2)$ are solved in the way described above. Because the sign of A is arbitrary, there are four combinations of baseline and rotation which are valid with respect to the original coplanarity equation: $(\mathbf{b}_1, R_1)$, $(\mathbf{b}_2, R_2)$, $(\mathbf{b}_1, R_2)$, and $(\mathbf{b}_2, R_1)$. However, only one of the four will be valid in practice. In order to determine this unique combination, we impose two constraints.
The first constraint is on the rotation angles. Because the two rotation matrices are symmetric, one of them corresponds to physically impractical rotation angles. In particular, α and β, as rotation angles about the x- and y-axis respectively, need to be in the range $[-\frac{\pi}{2}, \frac{\pi}{2}]$ for a reasonable stereo overlapping or vergence. However, to avoid using any subjective threshold, we can simply select one set of angles with

$$\min(\max(|\alpha_1|, |\beta_1|), \max(|\alpha_2|, |\beta_2|)) \tag{8.154}$$

After finding the unique rotation matrix R, we can now determine the unique baseline by using the second constraint, which requires the imaged objects and the image plane to be on the same side of the perspective centre. The unique set of baseline components can be determined in the following manner.

For any pair of homologous image points (x, y) and (x', y'), let

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} x - x_c \\ y - y_c \\ -f \end{pmatrix}, \quad \begin{pmatrix} u' \\ v' \\ w' \end{pmatrix} = R \begin{pmatrix} x' - x'_c \\ y' - y'_c \\ -f' \end{pmatrix} \tag{8.155}$$

Let $(X\,Y\,Z)^T$ denote the corresponding object point represented in the left camera coordinate system, then

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \kappa \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} b_x \\ b_y \\ b_z \end{pmatrix} + \kappa' \begin{pmatrix} u' \\ v' \\ w' \end{pmatrix} \tag{8.156}$$

where κ and κ' are two scale factors, which can be solved as

$$\begin{pmatrix} \kappa \\ \kappa' \end{pmatrix} = \frac{\begin{pmatrix} -v' & u' \\ -v & u \end{pmatrix} \begin{pmatrix} b_x \\ b_y \end{pmatrix}}{\begin{vmatrix} u & -u' \\ v & -v' \end{vmatrix}} = \frac{\begin{pmatrix} -w' & u' \\ -w & u \end{pmatrix} \begin{pmatrix} b_x \\ b_z \end{pmatrix}}{\begin{vmatrix} u & -u' \\ w & -w' \end{vmatrix}} = \frac{\begin{pmatrix} -w' & v' \\ -w & v \end{pmatrix} \begin{pmatrix} b_y \\ b_z \end{pmatrix}}{\begin{vmatrix} v & -v' \\ w & -w' \end{vmatrix}} \tag{8.157}$$

An appropriate equation from the above three equations may be selected to compute κ or κ'. For example, if both b_x and b_y are nonzero, and if $uv' - u'v \neq 0$, then

$$\kappa = \frac{u'b_y - v'b_x}{u'v - v'u} \tag{8.158}$$

Alternatively, to avoid singular cases, we can use all three equations (8.157) to solve for κ and κ'

$$\begin{pmatrix} \kappa \\ \kappa' \end{pmatrix} = (U^T U)^{-1} U^T \begin{pmatrix} b_x \\ b_y \\ b_z \end{pmatrix} \tag{8.159}$$

where

$$U = \begin{pmatrix} u & -u' \\ v & -v' \\ w & -w' \end{pmatrix} \tag{8.160}$$

With κ and κ' solved as above, we can now consider the Z coordinate of the object point. We require

$$Z = -\kappa f < 0 \tag{8.161}$$

therefore,

$$\kappa > 0 \tag{8.162}$$

With the unique rotation matrix R determined by using the first constraint, two symmetric baseline solutions lead to two κ's that are opposite in sign. The baseline with positive κ is then selected as the final correct solution.

8.6 An Iterative Least-Squares Solution for Explicit Parameters

After the general coplanarity matrix D is solved, the 7 explicit parameters: two principal distances f, f'; baseline components b_x, b_y, b_z; and rotation angles α, β, γ can be solved via a direct closed-form solution to be described later. We will therefore assume here that an initial approximation of these 7 parameters is available.

Since an approximate value of each baseline component is known, we may normalize the baseline vector by the maximum absolute value of the baseline components. That is, let $b_{\max} = \max(|b_x|, |b_y|, |b_z|)$ and normalize $\mathbf{b}$ as $(1/b_{\max})\mathbf{b}$. Without loss of generality, we consider the case where $|b_x| = b_{\max}$. This normalization effectively sets b_x to ± 1, and in all the subsequent iterations, only b_y and b_z will be updated. Note that this initialization of baseline components is, in essence, different from arbitrarily fixing one of the three totally unknown components to 1 while taking the other two as variables.

We now have a total of 7 parameters f, f', b_y, b_z, α, β, γ, together with their known initial approximations. Our subsequent task is to refine these approximations via an iterative least-squares solution. Let $\mathbf{k}$ be the unknown correction vector to these approximate values viz

$$\mathbf{k} = (\Delta f \quad \Delta f' \quad \Delta b_y \quad \Delta b_z \quad \Delta \alpha \quad \Delta \beta \quad \Delta \gamma)^T. \tag{8.163}$$

Our aim here is to use $\mathbf{k}$ to gradually update the 7 aforementioned parameters. The iteration process terminates when $\mathbf{k}$ becomes insignificant.

For each i-th pair of homologous image points, 4 corrections v_{x_i}, v_{y_i}, $v_{x'_i}$, and $v_{y'_i}$ are, respectively, associated with the four measurements x_i, y_i, x'_i, and y'_i. With n pairs of homologous image points, the vector of corrections to the homologous image point measurements can be described by the following $4n$-vector:

$$\mathbf{v} = (v_{x_1} \;\; v_{y_1} \;\; v_{x'_1} \;\; v_{y'_1} \;\; v_{x_2} \;\; v_{y_2} \;\; v_{x'_2} \;\; v_{y'_2} \;\; \ldots \;\; v_{x_n} \;\; v_{y_n} \;\; v_{x'_n} \;\; v_{y'_n})^T. \tag{8.164}$$

Since the special coplanarity matrix A is a function of five parameters b_y, b_z, α, β, γ (with b_x set to ± 1), the special coplanarity equation (8.11) can be written as

$$F(f, f', b_y, b_z, \alpha, \beta, \gamma) = \begin{pmatrix} x - x_c \\ y - y_c \\ -f \end{pmatrix}^T A(b_y, b_z, \alpha, \beta, \gamma) \begin{pmatrix} x' - x'_c \\ y' - y'_c \\ -f' \end{pmatrix} = 0 \tag{8.165}$$

Linearizing this function gives the following observation equation for n pairs of homologous image points:

$$G\,\mathbf{v} = H\,\mathbf{k} - L, \tag{8.166}$$

where

$$G = -\begin{pmatrix} F_{x,1} & F_{y,1} & F_{x',1} & F_{y',1} & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & F_{x,2} & F_{y,2} & F_{x',2} & F_{y',2} & \ldots & 0 & 0 & 0 & 0 \\ & & & & & & \vdots & & & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & F_{x,n} & F_{y,n} & F_{x',n} & F_{y',n} \end{pmatrix} \tag{8.167}$$

$$
\begin{aligned}
H &= \begin{pmatrix} F_{f,1} & F_{f',1} & F_{b_y,1} & F_{b_z,1} & F_{\alpha,1} & F_{\beta,1} & F_{\gamma,1} \\ F_{f,2} & F_{f',2} & F_{b_y,2} & F_{b_z,2} & F_{\alpha,1} & F_{\beta,1} & F_{\gamma,1} \\ & & & \vdots & & & \\ F_{f,n} & F_{f',n} & F_{b_y,n} & F_{b_z,n} & F_{\alpha,n} & F_{\beta,n} & F_{\gamma,n} \end{pmatrix} \qquad (8.168) \\
L &= -\begin{pmatrix} F_1 & F_2 & \dots & F_n \end{pmatrix}^T . \qquad (8.169)
\end{aligned}
$$

Here, $F_{m,i}$ denotes the partial derivative $F_m = \frac{\partial F}{\partial m}$ of the function F in (8.165) with respect to variable m computed for the i-th pair of homologous image points (x_i, y_i, x_i', y_i'); F_i denotes the value of function F computed for the i-th pair of points.
Under the least-squares criterion, we seek

$$
\min\{\, \mathbf{v}^T W \mathbf{v} - 2\lambda(G\mathbf{v} - H\mathbf{k} + L)\,\}, \qquad (8.170)
$$

where W is the $4n \times 4n$ weight matrix of the $4n$ image point measurements, λ is the Lagrangian multiplier. After some trivial derivation, the following solution for the correction vector $\mathbf{k}$ is achieved:

$$
\mathbf{k} = (H^T\,(G\,W^{-1}\,G^T)^{-1}\,H)^{-1}\,H^T\,(G\,W^{-1}\,G^T)^{-1}\,L \qquad (8.171)
$$

At each iteration, a new correction vector is computed as above and is used to refine the approximations of the 7 imaging parameters.
Given that good initial approximate values are obtained from the closed-form algebraic solutions [Pan et al, 1995; Pan 1997], this iterative linearized least-squares solution will converge to the precise values. The convergence of the iteration is confirmed from the many numerical tests on real images that have been accomplished.

8.7 Summary

Relative orientation of two stereo images, in the sense of traditional photogrammetry, refers to determining the baseline vector and the relative rotation matrix (or angles) of the two images, totally involving 5 free paramters, from a sufficient number of homologous image points. General relative orientation proposed by Pan (1995, 1997) is a generalization of this traditional definition to solving for 7 paramters including two principal distances (focal lengths) and the 5 relative orientation parameters. The relative geometry of two stereo images is captured in the coplanarity equations defined by homologous image points. The explicit coplanarity equations can be recast into an implicit form whose coefficients are grouped into a matrix, called coplanarity matrix. In principle, the elements of the coplanarity matrix can be determined through some closed-form solutions. However, the stability of the whole solution is mainly dependent on the precise solution of the coplanarity matrix. Therefore, two iterative non-linear least-square solutions involving singular value decomposition are presented to tackle this problem. Afterwards, two focal lengths are solvable in closed-form from the coplanarity matrix; and the baseline vector and rotation matrix and angles are determined also in closed-form. The total 7 explicit parameters can be fine-tuned through an iterative least-square solution using the original image measurements.

8.8 Excercises

1. Consider three overlapping images which form 3 stereo pairs. Suppose a sufficient number of homologous image points can be identified on all the three images. Derive a solution for the general relative orientation of the three images. Find out how many free paramters can be determined from pure image measurements.

2. Given a stereo head on which two stereo cameras are fixed. The relative orientation of two cameras is rigidly fixed when the stereo head moves. Consider two consecutive times t_0 and t_1. A stereo image pair is taken at each time. Derive a more robust solution to the general relative orientation of two cameras using the acquired two pairs of stereo images.

8.9 References

- Abdel-Aziz K. (1971): Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry. *Proc. ASP/UI Symposium on Close-Range Photogrammetry*, Urbana, Illinois, 1971, pp. 1-8.

- Brandstätter G. (1991): Zur relativen Orientierung projecktiver Bündel. Zeitschrift für Photogrammetrie und Fernerkundung, 59 Jahrgang, Heft 6, 199-212.

- Brandstätter G. (1996): On critical configurations of projective stereo correlation. Internatinal Archives of Photogrammetry and Remote Sensing (IAPRS), Vol. XXXI, B3, pp. 77-81, Vienna.

- Bruss A.R. and Horn B.K.P. (1983): Passive navigation. *Computer Vision, Graphics, and Image Processing* **21**(1): 3-20.

- Deriche R., Zhang Z., Luong Q.-T., and Faugeras O.D. (1994): Robust recovery of the epipolar geometry for an uncalibrated stereo rig. *Proc. ECCV'94*, pp. 567-576.

- Faugeras O.D. (1993): *Three-Dimensional Computer Vision: a geometric viewpoint.* MIT Press.

- Faugeras O.D., Lustman F. and Toscani G. (1987): Motion and structure from motion from point and line matches. *Proc. 1st ICCV*, 1987.

- Faugeras O.D., Luong Q.-T., and Maybank S.J. (1992): Camera self-calibration: Theory and experiments. *Proc. ECCV'92*, 1992, pp. 321-334.

- Finsterwalder S. (1899): Die Geometrischen Grundlagen der Photogrammetrie. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, Band VI, Teubner Press, Leipzig.

- Finsterwalder S. (1932): Die Hauptaufgabe der Photogrammetrie. *Bayer. Akad. der Wissehschaften München, Sit.-ber. der math.-phys.* Kl., pp. 115-131.

- Fourcade H.G. (1926): A new method of aerial surveying. *Transactions of the Royal Society of South Africa*, XIV(1):93-112, 1926. On page 49 of this issue, reproduced is a part of a letter of 5th June, 1903, written by Fourcade to A. Laussadat (a founder of photogrammetry), mentioning the theory of relative orientation. The theory was presented on 16th August, 1905 at a meeting of the British Association for the Advancement of Science.

- Hartley R.I. (1992): Estimation of relative camera positions for uncalibrated cameras. *Proc. ECCV'92*, pp. 579-587.

- Hartley R.I. (1995): On the defence of the eight-point algorithm. *Proc. ICCV'95*, pp. 1064-1070.

- Heisse O. (1863): Die cubische Gleichung, von welcher die Lölung des Problems der Homographies von M. Chasles abhängt. *Journal f. Reine und Angew. Math.* 62, 188-192.

- Kruppa E. (1913): Zur Ermittlung eines Objektes aus zwei Perspektiven mit inner Orientierung. Sitz.-Ber. Akad. Wiss., Wien, Math. Naturw. Kl (Abt. IIa. 122), 1939-1948.

- Horn1990 B.K.P. Horn, Relative orientation. *IJCV*, 4, 1990, 59-78

- Huang T.S. and Faugeras O.D. (1989): Some properties of the E matrix in two-view motion estimation. *IEEE T-PAMI* 11(12), 1310-1312.

- Khlebnikova T.A. (1983): Determining relative orientation angles of oblique aerial photographs. *Mapping Science and Remote Sensing*, 21(1), 95-100.

- Longuet-Higgins H.C. (1981): A computer algorithm for reconstructing a scene from two projections. *Nature*, 293(10): 133-135.

- Longuet-Higgins H.C. and Prazdney K. (1980): The interpretation of a moving retinal image, *Proc. Roy. Soc. London* B 208, pp. 385-397.

- Maybank S.J. and Faugeras O.D. (1992): A theory of self-calibration of a moving camera. *IJCV*, 8(2): 123-151.

- Niini I. (1993): Solution of the interior orientation without control points. The Photogrammetric Journal of Finland, 13(2):78-84.

- Niini I. (1994): Relative orientation of multiple images using projective singular correlation. *International Archives of Photogrammetry and Remote Sensing*, vol. 30, part 3/2, pp. 615-621, Munich.

- Pan H.-P. (1997): The Kernel of a Direct Closed-Form Solution to General Relative Orientation. *Proc. International Workshop on Image Analysis and Information Fusion* (IAIF'97), Adelaide, Australia, November.

- Pan H.-P., Brooks M.J. and Newsam G.N. (1995): Image resituation: initial theory. *Proc. Videometrics IV*, Philadelphia, USA, SPIE Proceedings 2598, pp. 162-173.

- Pan H.-P., Hyunh D.Q., and Hamlyn G.K. (1995): Two-image resituation: practical algorithm. *ibid*, pp. 174-190.

- Schmid H. (1954): An analytical treatment of the orientation of a photogrammetric camera. *Photogrammetric Engineering*, XX, 765-781.

- Schut G. (1955): Analytical aerial triangulation and comparison between it and instrumental aerial triangulation. *Photogrammetria*, XII, 1955/56, 311-318.

- Stefanovic P. (1973): Relative orientation - a new approach. *ITC Journal*, (ITC, Enschede, Netherlands), 417-448.

- Sturm R. (1969): Das Problem der Projektivität und seine Ansendung auf die Flächen zweiten Grades. Math. Ann., 1:533-574.

- Thompson E.H. (1959): A rational algebraic formulation of the problem of relative orientation. *Photogrammetric Record*, 3(14): 152-159.

- Torr P. (1995): Motion Segmentatin and Outlier Detection. PhD thesis, Department of Engineering Science, University of Oxford.

- Tsai R.Y. and Huang T.S. (1984): Uniqueness and estimation of 3-D motion parameters of rigid bodies with curved surfaces. *IEEE T-PAMI*, 6: 13-27.

- Wang Z.Z. Wang (1979): *Principles of Photogrammetry*, Surveying and Mapping Press, Beijing, China, 1979, (English version is available since 1990).

- Weng J., Huang T.S., and Ahuja N. (1993): *Motion and Structure From Image Sequences*, Springer-Verlag.

- Zhang Z., Deriche R., Faugeras O., and Luong Q.-T. (1995): A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. Artificial Intelligence, 78:87-119.