ELSEVIER
ISPRS Journal of Photogrammetry & Remote Sensing 51 (1996) 299–307
PHOTOGRAMMETRY & REMOTE SENSING

An algorithm for object reconstruction without interior orientation

J. Shan

Department of Geodetic Science, Stuttgart University, Keplerstr. 11, D-70174 Stuttgart, Germany

Received 28 January 1996; accepted 24 April 1996

Abstract

This paper presents a novel algorithm for object reconstruction without interior orientation. First we introduce the Thompson and Longuet-Higgins equation as well as the fundamental matrix. After defining the affine model, we show that some of its components can be linearly withdrawn from the fundamental matrix, which in turn is linearly determined up to a scale factor by minimal eight image correspondences. The object reconstruction is then performed linearly via a 3D affine transformation. Unlike the well-known DLT algorithm where minimum six known points are required on each image of a stereopair, our algorithm requires that only four of them appear on the second image. In addition to an accuracy fully compatible with the DLT algorithm, tests with an aerial stereopair also show the robustness of our algorithm, both to the configuration of known points and to the image deformation.

1. Motivations

For quite a long time it seemed to be a rule for photogrammetry that the interior orientation has to be completed, or equivalently, the interior elements of the camera must be known, before any other photogrammetric computation is done. This was overthrown by the time the well-known DLT (Direct Linear Transformation) algorithm was published by Abdel-Aziz and Karara in 1971 (cf., Slama, 1980, pp. 801–803). It directly relates the object point to its image coordinates which could be measured in any oblique coordinate system without knowledge of interior elements of the camera. This problem seems to be fully solved if we neglect the drawbacks of the DLT algorithm. In fact it recovers an object directly from its images rather than from the photogrammetric model, therefore, the inherent information behind the photogrammetric stereopair is not fully utilized. Moreover, it requires at least *six* known points on *each* of the images to reconstruct the object. Keeping those issues in mind, the question arises: how could we reconstruct an object without knowledge of interior orientation by fully employing the photogrammetric information behind a stereopair? Our motivation also has its deep roots in computer vision and close-range photogrammetry where uncalibrated camera is widely adopted and the interior elements are either unknown or different from image to image.

Besides those theoretical considerations, our motivation also has its practical background. Although there seem to be no theoretical impossibilities to search and localize the fiducial marks in the digital image fully automatically, each of current digital photogrammetric systems performs this interactively or semi-automatically, probably due to the large searching area in huge image data and the various types of fiducial marks. Therefore, if the interior orientation procedure becomes not inevitable anymore, we could achieve progress towards the automation of digital photogrammetry.

Linear solution is always very beneficial, es-

0924-2716/96/$15.00
PII S0924-2716(96)00024-X

pecially in computer vision and close-range photogrammetry, where finding reasonable initial values is crucial to the success of iterative algorithms.

Regarding these backgrounds, our purpose is to find a linear solution for object reconstruction without interior orientation, which can fully use the information within a stereopair and has less requirements on known points than the DLT algorithm.

2. Review of related work and our scope

As our topic falls in both photogrammetry and computer vision, the related work in both areas should be mentioned. Photogrammetrists seem to rely much on the DLT algorithm and hence on a sufficient number of known object points, when the interior elements are unknown. So far, probably no algorithm is widely acknowledged to take advantage of the photogrammetric model. In contrast, besides being interested in camera calibration, scientists in the computer vision area have fully studied the problem "motion or relative displacement estimation from uncalibrated camera". Most recently, quite a few papers have been focused on this issue. They claimed that without interior elements the object can be reconstructed up to either an affine or a perspective transformation, if only eight image correspondences are given (Faugeras, 1992; Hartley et al., 1992; Hartley, 1992). Obviously this is of fundamental importance for object reconstruction without interior orientation (Faugeras, 1993; Hartley and Mundy, 1993).

A linear solution for model reconstruction may date back to the well-known contribution of Longuet-Higgins (1981) in the realm of computer vision. However, the basic idea behind this solution essentially originated from the early work of Thompson (1968), where he expressed the coplanarity equation via an unknown 3 × 3 matrix which is acknowledged today as *essential matrix* (cf., Longuet-Higgins, 1981; Huang and Faugeras, 1989; Faugeras and Maybank, 1990; Hartley, 1992; Hartley and Mundy, 1993). It was not until recently that photogrammetrists recalled Thompson's idea. Brandstätter (1992) employed this idea for image rectification. The work of Wang (1995) threw a light on this idea upon which a linear algorithm was designed to reconstruct the photogrammetric model with the aid of the interior orientation. Most recently, the stability of this algorithm was studied by Deriche et al. (1994), Förstner (1995) and Luong and Faugeras (1994).

Our research is highly inspired by the work of Hartley et al., Faugeras et al. and Wang. In Section 3 we start from an affine transformation in image space in terms of homogeneous coordinates, and then generalize the Thompson and Longuet-Higgins equation to the case of unknown interior elements. Section 4 is focused on the affine model and the recovery of its components. Unlike Faugeras' work (Faugeras, 1992), where traditional projective geometry is utilized, we fully take advantage of the properties of the skew-symmetric matrix (cf., Appendix A) and make our development as parallel as possible to photogrammetry both in concept and in the form of formulae. After defining the affine model parallel to the traditional one, we show that some of its components can be withdrawn from the so-called *fundamental matrix*. This leads to a complete employment of a stereopair. In Section 5 we use a 3D affine transformation to fully recover the affine model as well as to orient it to the object frame. Unlike the well-known DLT algorithm, where a minimum of six known points are required on each image of a stereopair, our algorithm allows that one image may have only four of them. Moreover, this solution is fully linear. Its algorithmic realization is described in Section 6. Tests with an aerial stereopair show that our algorithm is robust both to the configuration of known points and to the image deformation. Results fully compatible with the DLT algorithm are obtained as well.

3. Thompson and Longuet-Higgins equation

In this section we derive the Thompson and Longuet-Higgins equation, which plays a fundamental role in our problem. A short comment is thereafter made on the fundamental matrix.

Referring to Fig. 1 where object point P is imaged on a stereopair, we have the well-known coplanarity equation (Slama, 1980, pp. 54–56)

$$\boldsymbol{x}_1^{\mathrm{T}}[\boldsymbol{b} * (R\boldsymbol{x}_2)] = 0. \tag{1}$$

In Eq. (1)

$$\boldsymbol{b} = (B_X \quad B_Y \quad B_Z)^{\mathrm{T}} \tag{2}$$

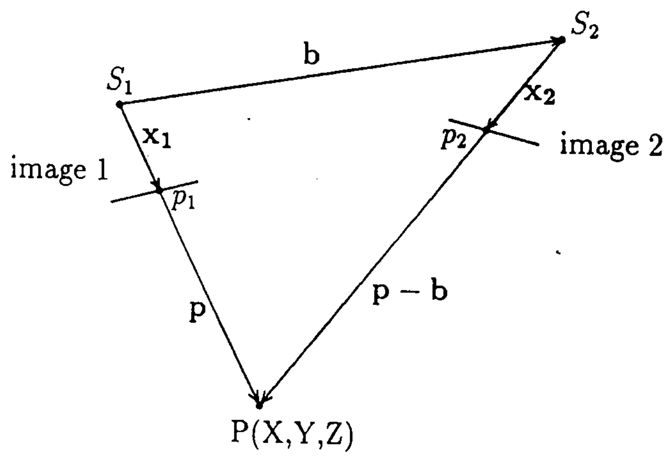

Fig. 1. Geometry of a stereopair.

is the base component vector, R is the orthogonal rotation matrix of the second image relative to the first one, which is assumed to be a reference, and

$$\boldsymbol{x}_1 = (x_1 \quad y_1 \quad -f_1)^{\mathrm{T}}, \quad \boldsymbol{x}_2 = (x_2 \quad y_2 \quad -f_2)^{\mathrm{T}} \tag{3}$$

are coordinates of conjugate image points p_1, p_2 in their image spaces, respectively. In Eq. (1) $*$ denotes the cross product of two vectors.

For any vector $\boldsymbol{x}$ we have (cf., Eq. (A.3))

$$\boldsymbol{b} * \boldsymbol{x} = B\boldsymbol{x}, \tag{4}$$

where B is a 3×3 skew-symmetric matrix, the entries of which are composed of the elements of $\boldsymbol{b}$, i.e.,

$$B = \begin{pmatrix} 0 & -B_Z & B_Y \\ B_Z & 0 & -B_X \\ -B_Y & B_X & 0 \end{pmatrix} \tag{5}$$

Applying Eq. (4) to Eq. (1) yields

$$\boldsymbol{x}_1^{\mathrm{T}} E \boldsymbol{x}_2 = 0, \tag{6}$$

where

$$E = BR \tag{7}$$

Eq. (6) is namely the *Thompson and Longuet-Higgins equation*, which was initially derived by Thompson (1968) and rediscovered by Longuet-Higgins (1981). Matrix E, which is the product of the *base component matrix* B and the orthogonal rotation matrix R, is named as *essential matrix* by Longuet-Higgins (1981) and thereafter widely accepted and studied in the realm of computer vision (Huang and Faugeras, 1989; Faugeras and Maybank, 1990). Note that Eq. (6) is only valid for the case when interior orientation is performed, as (x_1, y_1) and (x_2, y_2) are originated at their corresponding principal points.

It is quite direct to generalize Eq. (6) to the case when the interior orientation is not done. Suppose image points are measured in an arbitrary oblique coordinate system $(\bar{x}, \bar{y})$ which is considered as a linear or an affine transformation of (x, y), namely we have

$$\boldsymbol{x}_1 = A_1 \bar{\boldsymbol{x}}_1, \quad \boldsymbol{x}_2 = A_2 \bar{\boldsymbol{x}}_2, \tag{8}$$

where

$$A_1 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & -f_1 \end{pmatrix}_1,$$

$$A_2 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & -f_2 \end{pmatrix}_2, \tag{9}$$

$$\bar{\boldsymbol{x}}_1 = (\bar{x}_1 \quad \bar{y}_1 \quad 1)^{\mathrm{T}}, \quad \bar{\boldsymbol{x}}_2 = (\bar{x}_2 \quad \bar{y}_2 \quad 1)^{\mathrm{T}}. \tag{10}$$

$\bar{\boldsymbol{x}}_1$ and $\bar{\boldsymbol{x}}_2$ are essentially the homogeneous coordinates of image points. Substituting Eq. (8) back into Eq. (6) we obtain

$$\bar{\boldsymbol{x}}_1^{\mathrm{T}} \bar{E} \bar{\boldsymbol{x}}_2 = 0, \tag{11}$$

where

$$\bar{E} = A_1^{\mathrm{T}} E A_2 = A_1^{\mathrm{T}} B R A_2 \tag{12}$$

is called *fundamental matrix* (Faugeras et al., 1992). Eqs. (11) and (12) are the generalization of Eqs. (6) and (7), and elementary for performing relative orientation and recovering components of a stereopair when interior elements are unknown. The following analyses are essential for the implementation of Eq. (11).

(a) Eq. (11) is linear with respect to the entries of matrix $\bar{E}$. It can only be determined up to a scale factor. This fact follows if we notice that Eq. (11) is homogeneous and therefore multiplication with a scale factor does not change its value. Hence there are only *eight linearly independent* parameters in matrix $\bar{E}$. Therefore we need a minimum of eight image correspondences to linearly determine the $\bar{E}$ matrix.

(b) There are only *seven degrees of freedom* among the eight linearly independent parameters as they meet $|\bar{E}| = 0$ (since $|B| = 0$, cf., Eq. (A.2)). This is coherent with Sturm's algorithm of computing epipoles (cf., Faugeras et al., 1992; Maybank and Faugeras, 1992), which is known as Chasles' homograph problem in mathematics (cf., Buchanan, 1993). However, due to its non-linearity, the solution is not unique.

(c) There are three degrees of rank deficiency in the reconstruction of the affine model with pure image correspondences. A simple numeric counting may show this fact. From the 2×11 degrees of perspective transformation in a stereopair, 7 could be recovered in the relative orientation step and 12 in the absolute orientation step. There remain $2 \times 11 - (7 + 12) = 3$ degrees of deficiency to be removed by choosing a basis on the image plane, or by leaving them to the exterior orientation step, which will yield $12 + 3 = 15$ exterior parameters.

4. Affine model and recovery of its components

In the following derivation we assume that the fundamental matrix $\bar{E}$ has been determined up to a scale factor by at least eight image correspondences using Eq. (11).

4.1. Affine model and its components

Referring to Fig. 1 and applying the collinear condition for the first and second image, respectively, we obtain

$$\boldsymbol{p} = \lambda_1 \boldsymbol{x}_1, \quad \boldsymbol{p} = \lambda_2 R \boldsymbol{x}_2 + \boldsymbol{b}, \tag{13}$$

where

$$\boldsymbol{p} = (X \quad Y \quad Z)^{\mathrm{T}} \tag{14}$$

gives the coordinates of point P in the first image space coordinate system, λ_1 and λ_2 are scale factors of vectors $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ relative to vectors $\boldsymbol{p}$ and $\boldsymbol{p} - \boldsymbol{b}$. Substituting Eq. (8) into Eq. (13) yields

$$\boldsymbol{p} = \lambda_1 A_1 \bar{\boldsymbol{x}}_1, \quad \boldsymbol{p} = \lambda_2 R A_2 \bar{\boldsymbol{x}}_2 + \boldsymbol{b}. \tag{15}$$

The next step is essential for solving our problem. We write Eq. (15) in a way similar to Eq. (13):

$$\bar{\boldsymbol{p}} = \bar{\lambda}_1 \bar{\boldsymbol{x}}_1, \quad \bar{\boldsymbol{p}} = \bar{\lambda}_2 \bar{R} \bar{\boldsymbol{x}}_2 + \bar{\boldsymbol{b}}, \tag{16}$$

where

$$\bar{\boldsymbol{p}} = A_1^{-1} \boldsymbol{p} = (\bar{X} \quad \bar{Y} \quad \bar{Z})^{\mathrm{T}}, \tag{17}$$

$$\bar{\boldsymbol{b}} = A_1^{-1} \boldsymbol{b} = (\bar{B}_X \quad \bar{B}_Y \quad \bar{B}_Z)^{\mathrm{T}}, \tag{18}$$

$$\bar{R} = A_1^{-1} R A_2. \tag{19}$$

$\bar{\lambda}_1$ and $\bar{\lambda}_2$ are proportional to λ_1 and λ_2, respectively, as they may take into account the multiplication factor inherent in the $\bar{E}$ matrix.

We define $\bar{\boldsymbol{p}} = (\bar{X} \quad \bar{Y} \quad \bar{Z})^{\mathrm{T}}$ as the *affine coordinates* of point P, as it is a linear, i.e., affine transformation of its Cartesian coordinates. The collection of all affine points forms an affine model of the object. Similarly, $\bar{\boldsymbol{b}}$ is known as the *affine base component vector*, and $\bar{R}$ is specified as the *affine rotation matrix*.

Now we are due to rewrite the fundamental matrix $\bar{E}$ in Eq. (12) as

$$\bar{E} = (A_1^{\mathrm{T}} B A_1)(A_1^{-1} R A_2) = \bar{T} \bar{R}, \tag{20}$$

where $\bar{T} = A_1^{\mathrm{T}} B A_1$ is a skew-symmetric matrix (cf. Eq. (A.6). As $\bar{T}\bar{\boldsymbol{b}} = \boldsymbol{0}$, we are led immediately to (cf., Eq. (A.9))

$$\frac{\bar{T}_X}{\bar{B}_X} = \frac{\bar{T}_Y}{\bar{B}_Y} = \frac{\bar{T}_Z}{\bar{B}_Z} = c,$$

i.e., $\bar{T}$ is proportional to $\bar{B}$. As any multiplication factor in matrix $\bar{T}$ can be taken into account within the fundamental matrix, we may simple let $c = 1$. Thus, Eq. (20) becomes

$$\bar{E} = \bar{B} \bar{R}, \tag{21}$$

where

$$\bar{B} = \begin{pmatrix} 0 & -\bar{B}_Z & \bar{B}_Y \\ \bar{B}_Z & 0 & -\bar{B}_X \\ -\bar{B}_Y & \bar{B}_X & 0 \end{pmatrix} \tag{22}$$

is defined as the *affine base component matrix*, analogous to B in Eq. (5).

Being parallel to the essential matrix E of Eq. (7), Eq. (21) reveals the following basic and important fact:

The fundamental matrix $\bar{E}$ can be decomposed as a product of $\bar{B}$ and $\bar{R}$, where $\bar{B}$ is a skew-symmetric *matrix composed of the affine base components, representing the displacement of the second perspective centre in the first affine image space, and $\bar{R}$ is an affine rotation matrix representing the orientation of the second camera with respect to the first.*

4.2. Recovery of the affine model components

We start with computing the affine base component vector $\bar{\boldsymbol{b}}$. Noticing that $B^{\mathrm{T}}\boldsymbol{b} = \mathbf{0}$ (cf., Eq. (A.5)), we obtain

$$\bar{E}^{\mathrm{T}}\bar{\boldsymbol{b}} = \mathbf{0}. \tag{23}$$

This is a set of homogeneous equations with $\bar{\boldsymbol{b}}$ as unknown and $\bar{E}^{\mathrm{T}}$ as known coefficients. As $|\bar{E}^{\mathrm{T}}| = |\bar{E}| = 0$, which means that Eq. (23) has non-zero solutions, by setting $\bar{B}_X$ equal to an arbitrary positive constant, we then could obtain the other two affine base components $\bar{B}_Y$ and $\bar{B}_Z$.

Once $\bar{B}$ is obtained via Eq. (23), we may move on to computing the matrix $\bar{R}$. However, as is pointed out in Section 3, there are three degrees of rank deficiency. This can also be seen from Eq. (21), which is now written column-wise as

$$\bar{B}\bar{r}_i = \bar{e}_i \quad (i = 1, 2, 3), \tag{24}$$

where $\bar{r}_i$ and $\bar{e}_i$ are the column component vectors of $\bar{R}$ and $\bar{E}$, respectively. Since $\operatorname{rank}(\bar{B}) = 2$, for each column of $\bar{R}$ we can only determine two parameters, namely *three* independent parameters (degrees of rank deficiency) remain altogether, which could not be *uniquely* determined within a stereopair. By choosing $\bar{r}_{11}$, $\bar{r}_{12}$, $\bar{r}_{13}$ as independent parameters, a way the solution will always be ensured, we have

$$\begin{aligned} \bar{r}_{21} &= (\bar{B}_Y\bar{r}_{11} + \bar{e}_{31})/\bar{B}_X & \bar{r}_{31} &= (\bar{B}_Z\bar{r}_{11} - \bar{e}_{21})/\bar{B}_X \\ \bar{r}_{22} &= (\bar{B}_Y\bar{r}_{12} + \bar{e}_{32})/\bar{B}_X & \bar{r}_{32} &= (\bar{B}_Z\bar{r}_{12} - \bar{e}_{22})/\bar{B}_X \\ r_{23} &= (\bar{B}_Y\bar{r}_{13} + \bar{e}_{33})/\bar{B}_X & \bar{r}_{33} &= (\bar{B}_Z\bar{r}_{13} - \bar{e}_{23})/\bar{B}_X \end{aligned} \tag{25}$$

Moreover, we can also obtain the length ratio, i.e., $\bar{\lambda}_2/\bar{\lambda}_1$, of the two conjugate projective rays. Equalizing the two equations in Eq. (16) and multiplying with $\bar{B}$ on both sides yields its least-squares solution:

$$\bar{k} = \frac{\bar{\lambda}_2}{\bar{\lambda}_1} = \frac{(\bar{E}\bar{\boldsymbol{x}}_2)^{\mathrm{T}}(\bar{B}\bar{\boldsymbol{x}}_1)}{(\bar{E}\bar{\boldsymbol{x}}_2)^{\mathrm{T}}(\bar{E}\bar{\boldsymbol{x}}_2)} \tag{26}$$

In summary, with pure image correspondences or the fundamental matrix, we could recover the *two ratios* of the three affine base components, as well as *six relationships* among the nine components of the affine rotation matrix. That amounts to altogether eight parameters corresponding to the linearly independent entries in the fundamental matrix. Moreover, for each image correspondence, we could determine its length ratio of the two conjugate projective rays. Since there are three degrees of rank deficiency, the affine model can not be fully reconstructed without extra object information.

Once $\bar{B}$ and $\bar{R}$ are determined (cf. next section), we are due to compute the coordinates of the affine model points. From Eq. (16) we have the least-squares solution of $\bar{\lambda}_1$:

$$\bar{\lambda}_1 = \frac{(\bar{\boldsymbol{x}}_1 - \bar{k}\bar{R}\bar{\boldsymbol{x}}_2)^{\mathrm{T}}\bar{\boldsymbol{b}}}{(\bar{\boldsymbol{x}}_1 - \bar{k}\bar{R}\bar{\boldsymbol{x}}_2)^{\mathrm{T}}(\bar{\boldsymbol{x}}_1 - \bar{k}\bar{R}\bar{\boldsymbol{x}}_2)}. \tag{27}$$

Finally, the model point is taken average of the two projective equations

$$\bar{\boldsymbol{p}} = \tfrac{1}{2}(\bar{\lambda}_1\bar{\boldsymbol{x}}_1 + \bar{\lambda}_2\bar{R}\bar{\boldsymbol{x}}_2 + \bar{\boldsymbol{b}}). \tag{28}$$

There is no doubt that in the above recovery procedure the key issue is to determine the $\bar{R}$ matrix, or equivalently, to determine its three independent parameters. To remove the rank deficiency, some authors have suggested to choose a basis of three points on the image plane (Faugeras, 1992), or to accomplish singular value decomposition of the fundamental matrix (Hartley, 1992; Hartley et al., 1992). However, these solutions are not unique. Our tests show that results from different bases may vary up to meters in object space. In this sense we cannot prefer this strategy. Instead, the complete recovery of $\bar{R}$ is included in the exterior orientation step, i.e., based on known object points.

5. Linear object reconstruction

In this section we first present the transformation between the object space and the affine model. Then a linear algorithm is designed to perform the exterior orientation of the *partially* recovered affine model.

It is trivial to show that the transformation between the object space and the affine model takes the form (affine transformation)

$$\bar{p} = Au = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \end{pmatrix} \begin{pmatrix} U \\ V \\ W \\ 1 \end{pmatrix}, \tag{29}$$

where $u^{\mathrm{T}} = (U \quad V \quad W \quad 1)$ gives the homogeneous coordinates of a point in object space, A is the transformation matrix, specified by twelve independent parameters.

Inserting Eq. (16) into Eq. (29) yields

$$Au = \bar{\lambda}_1 \bar{x}_1, \qquad Au = \bar{\lambda}_2 \bar{R} \bar{x}_2 + \bar{b}, \tag{30}$$

which has 12 (from A) + 3 (from $\bar{R}$) = 15 independent orientation parameters.

To design a linear solution, we eliminate $\bar{\lambda}_1$ in Eq. (30) and get two DLT-type equations

$$\bar{x}_1 = \frac{a_1 U + a_2 V + a_3 W + a_4}{c_1 U + c_2 V + c_3 W + c_4},$$
$$\bar{y}_1 = \frac{b_1 U + b_2 V + b_3 W + b_4}{c_1 U + c_2 V + c_3 W + c_4}. \tag{31}$$

The twelve parameters could be linearly determined up to a scale factor with six given object points appearing in the first image, namely the ratios $a_i' = a_i/c_4$, $b_i' = b_i/c_4$, $c_i' = c_i/c_4$ ($c_4' = 1$) are obtained.

Immediately after that, the remaining four parameters are determined linearly with the second set of Eq. (30) by minimal *four conjugate* given object points, i.e.,

$$A'u = \bar{k}\bar{\lambda}_1' \bar{R} \bar{x}_2 + \frac{1}{c_4}\bar{b}, \tag{32}$$

where matrix A' is composed of a_i', b_i', c_i' similar to matrix A,

$$\bar{\lambda}_1' = c_1' U + c_2' V + c_3' W + 1, \tag{33}$$

and $\bar{k}$ is computed from Eq. (26).

The object reconstruction is then finally performed by reversing Eq. (29)

$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}^{-1} \begin{pmatrix} \bar{X} - a_4 \\ \bar{Y} - b_4 \\ \bar{Z} - c_4 \end{pmatrix}. \tag{34}$$

Compared to the DLT algorithm, which relates every image independently to the object space, our algorithm fully employs the information within a stereopair. Therefore, only 15 instead of 2×11 parameters are dependent on known object points. To ensure a linear solution of this algorithm, one image of a stereopair should have a minimum of six known object points, and minimally four of them appear on the other one. It is apparent that in this minimum configuration of known points, the DLT algorithm fails, while a linear solution is available in our algorithm.

6. Algorithmic implementation and tests

In this section we first describe the implementation of our algorithm and then report and analyze the test results with an aerial stereopair.

6.1. Algorithmic implementation

The realization of our algorithm is composed of following steps.

Step 1: determining the fundamental matrix $\bar{E}$. We rewrite Eq. (11) in an operational way

$$(\bar{x}_2^{\mathrm{T}} \otimes \bar{x}_1^{\mathrm{T}}) \cdot \operatorname{vec} \bar{E} = 0, \tag{35}$$

where $\otimes$ means the Kronecker product and *vec* refers to the columnwise vectorization of a matrix.

As only the ratios among the entries of $\bar{E}$ could be determined, we may simply let one of its components be equal to one. It is proper to set $\bar{e}_{32} = 1$ as it is approximately equal to $\bar{B}_X \bar{r}_{22}$, which may never be zero. This setting also leads the right item in Eq. (35) to $-\bar{y}_1$, which is analogous to the y-parallax in the traditional sense. We include all image correspondences to determine $\bar{E}$.

Moreover, since there are only seven degrees of freedom in the fundamental matrix, the condition

$$|\bar{E}| = 0 \tag{36}$$

may also be included in the solution procedure with a properly chosen weight. However, unlike the experience of Barakat et al. (1994), our tests show that this condition can only stabilize the solution to a limited extent. Probably due to the large amount of image correspondences, the determination of the fundamental matrix is well-conditioned.

Table 1
Photographic parameters

Flight height:	ca. 2250 m
focal length:	88.94 mm
frame size:	230 mm × 230 mm
camera:	RC-10
overlap:	ca. 65%

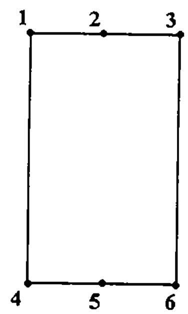

Fig. 2. GCPs distribution in the stereopair.

Step 2: determining the affine base component vector $\bar{b}$ or the affine base component matrix $\bar{B}$ with Eq. (23). In this step we set $\bar{B}_X$ equal to the image base length.

Step 3: computing the projective length ratio for each image correspondence with Eq. (26).

Step 4: determining the entries of matrix A' with six known points using Eq. (31).

Step 5: determining the three independent parameters in matrix $\bar{R}$ and the scale factor c_4 with Eq. (32).

Step 6: fully reconstructing the affine model with Eqs. (26), (27) and (28).

Step 7: reconstructing the object with Eq. (34) for each affine model point.

6.2. Tests and analyses

We use an aerial image stereopair to evaluate our algorithm. Its primary parameters and the distribution of the six ground control points (GCPs) are shown in Table 1 and Fig. 2, respectively.

Altogether 36 image points as well as their 3D ground coordinates are measured with an analytical plotter. The latter are treated as "best values" to check the validity of our algorithm. Moreover, the DLT algorithm and the traditional collinear algorithm are also implemented. In order to check the efficiency of our algorithm, results under different control configurations and various image deformations are presented respectively in Tables 2 and 3, in which all numerics are compared with the "best values".

Table 2
Results under different GCPs (in meters)

GCPs config. on the second image	RMSE to best values		
	σ_X	σ_Y	σ_Z
a: 1–2–3–4–5–6	1.936	1.595	1.722
b: 2–3–4–6	1.892	1.455	1.722
c: 1–2–4–5	1.944	1.648	1.723
d: 2–3–5–6	1.887	1.444	1.722
e: 2–4–5–6	1.915	1.539	1.722
f: 1–2–3–5	1.954	1.633	1.722
DLT algorithm	1.954	1.580	1.736
Colli. algorithm	1.376	1.365	1.745

Table 3
Results under image deformations (in meters)

Amount of image deform. parameters	RMSE to best values		
	σ_X	σ_Y	σ_Z
1. no deformation	1.892	1.455	1.722
2. $s = 1.1$, $\alpha = 10°$, $d = 10$ mm	1.926	1.460	1.715
3. $s = 0.9$, $\alpha = 20°$, $d = 20$ mm	1.907	1.466	1.719
4. $s = 1.3$, $\alpha = 30°$, $d = 30$ mm	1.923	1.465	1.748
5. $s = 0.7$, $\alpha = 40°$, $d = 40$ mm	1.911	1.488	1.743
DLT algorithm	1.954	1.580	1.736

In Table 2, the DLT algorithm and collinear algorithm are implemented with all six conjugate GCPs, while our algorithm is evaluated with different GCPs configuration on the second image and six common GCPs on the first image.

It is no wonder that the collinear algorithm holds the best results. The items a and DLT in Table 2 show that our algorithm obtains essentially the same rigorous results as the DLT algorithm when they have the same GCPs configuration. Through items b to f of Table 2, where the DLT algorithm is not accessible, our algorithm behaves completely robust to various GCPs configurations. The small differences, rarely up to maximally decimeters, are within the precision tolerance of GCPs themselves. Moreover, the most encouraging is that in each minimum GCPs configuration we could still reach the same accuracy as for the full GCPs configuration, a benefit due to the complete employment of the information within a stereopair.

Table 3 shows the results of our algorithm under different affine image deformations, where s, α and d refer to the scale factor, rotation angle and the disparity of the principal point, respectively. In order to testify the validity of our algorithm, simulated affine deformations based on these parameters are added on the original image observations, where the first and second image take different signs of the parameters, respectively. The GCPs configuration of this table is item b in Table 2. Since the DLT algorithm presents the same result under different image deformations, it is appended only in one row in Table 3. It is clearly seen that our algorithm is practically robust to different amounts of affine image deformations, since only trivial changes (maximum up to centimeters) might occur among them.

7. Conclusions

Object reconstruction without interior orientation can be linearly accomplished with the aid of the affine model. By making full use of the information in a stereopair we can determine 2 ratios of the affine base components and 6 relationships among the 9 entries of the affine rotation matrix. The partially reconstructed affine model is oriented to an object frame via determining 15 independent parameters. Unlike the DLT algorithm, where minimum 6 known points are required on each image of a stereopair, our algorithm allows one image to have only 4 of them. In addition to its completely compatible accuracy with the DLT algorithm, it is robust to control configurations and image deformations. This algorithm will find wide uses in object reconstruction in computer vision and photogrammetry with uncalibrated cameras or without interior orientation.

Acknowledgements

Prof. Dr.-Ing. E. Grafarend is gratefully thanked since he led the author to projective geometry and its application in computer vision.

Appendix A. Properties of a 3 × 3 skew-symmetric matrix

1. A 3×3 skew-symmetric matrix S is defined as

$$S = -S^{\mathrm{T}} = \begin{pmatrix} 0 & -S_Z & S_Y \\ S_Z & 0 & -S_X \\ -S_Y & S_X & 0 \end{pmatrix}. \tag{A.1}$$

It has only three independent non-zero entries and is singular, i.e.

$$|S| = 0. \tag{A.2}$$

2. For a 3×1 vector $\boldsymbol{x} = (X \quad Y \quad Z)^{\mathrm{T}}$, we have

$$\boldsymbol{s} * \boldsymbol{x} = S\boldsymbol{x}, \tag{A.3}$$

where $*$ denotes the cross product of two vectors, and

$$\boldsymbol{s} = (S_X \quad S_Y \quad S_X)^{\mathrm{T}} \tag{A.4}$$

is specified by the entries of matrix S. This can be easily showed, as

$$\boldsymbol{s} * \boldsymbol{x} = \begin{pmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ S_X & S_Y & S_Z \\ X & Y & Z \end{pmatrix} =$$

$$\begin{vmatrix} S_Y & S_Z \\ Y & Z \end{vmatrix} \boldsymbol{i} + \begin{vmatrix} S_Z & S_X \\ Z & X \end{vmatrix} \boldsymbol{j} + \begin{vmatrix} S_X & S_Y \\ X & Y \end{vmatrix} \boldsymbol{k} = S\boldsymbol{x},$$

where $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ are unit vectors in the directions of coordinate axes.

3. If $\boldsymbol{s}$ is composed of the entries of S, then

$$S\boldsymbol{s} = -S^{\mathrm{T}}\boldsymbol{s} = \boldsymbol{0}. \tag{A.5}$$

4. For any 3×3 matrix A, $\bar{S} = A^{\mathrm{T}}SA$ is also a skew-symmetric matrix, as

$$\bar{S}^{\mathrm{T}} = (A^{\mathrm{T}}SA)^{\mathrm{T}} = -A^{\mathrm{T}}SA = -\bar{S}. \tag{A.6}$$

5. It is easy to show

$$S\boldsymbol{x} = X\boldsymbol{s}, \tag{A.7}$$

where X is the skew-symmetric matrix specified by the components of $\boldsymbol{x}$ according to Eq. (A.1).

6. The solution of the homogeneous equation system

$$S\boldsymbol{x} = \boldsymbol{0} \tag{A.8}$$

is

$$\frac{X}{S_X} = \frac{Y}{S_Y} = \frac{Z}{S_Z}. \tag{A.9}$$

References

Barakat, H., Doucette and P., Mikhail, E.M., 1994. Photogrammetric analysis of image invariance. Spatial Information from Digital Photogrammetry and Computer Vision, International Archives of Photogrammetry and Remote Sensing, Vol. 30, Part 3/1, pp. 25–34.

Brandstätter, G., 1992. Notes on the direct projective transformation of general stereo pairs into the rigorous normal case by image correlation. International Archives of Photogrammetry and Remote Sensing, Vol. 29, Part B3, pp. 701–706.

Buchanan, T., 1993. Photogrammetry and projective geometry – a historical survey. In: E.B. Barrett and D.M. McKeown (Editors), Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision, Vol. 1944. SPIE, pp. 82–91.

Deriche, R., Zhang, Z., Luong, Q.-T. and Faugeras, O.D., 1994. Robust recovery of the epipolar geometry for an uncalibrated stereo rig. In: J.-O. Eklundh (Editor), Computer Vision-ECCV'94, Lecture Notes in Computer Science, No. 800. Springer Verlag, pp. 567–576.

Faugeras, O.D., 1992. What can be seen in three dimensions with an uncalibrated stereo rig? In: G. Sandini (Editor), Computer Vision-ECCV'92, Lecture Notes in Computer Science, No. 588. Springer Verlag, pp. 563–578.

Faugeras, O.D., 1993. Computer vision research at INRIA. Int. J. Computer Vision, 10(2): 91–99.

Faugeras, O.D. and Maybank, S., 1990. Motion from point matches: multiplicity of solutions. Int. J. Computer Vision, (4): 225–246.

Faugeras, O.D., Luong, Q.-T. and Maybank, S.J., 1992. Camera self-calibration: theory and experimences. In: G. Sandini (Editor), Computer Vision-ECCV'92, Lecture Notes in Computer Science, No. 588. Springer Verlag, pp. 321–334.

Förstner, W., 1995. Generic estimation procedures for orientation with minimum and redundant information. Proc. Workshop on Camera Calibration and Orientation, Washington.

Hartley, R.I., 1992. Estimation of relative camera positions for uncalibrated cameras. In: G. Sandini (Editor), Computer Vision-ECCV'92, Lecture Notes in Computer Science, No. 588. Springer Verlag, pp. 579–587.

Hartley, R.I. and Mundy, J.L., 1993. The relationship between photogrammetry and computer vision. In: E.B. Barrett and D.M. McKeown (Editors), Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision. Vol. 1944, SPIE, pp. 92–105.

Hartley, R.I., Gupta, R. and Chang, T., 1992. Stereo from uncalibrated cameras. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp. 761–764.

Huang, T.S. and Faugeras, O.D., 1989. Some properties of the E matrix in two-view motion estimation. IEEE Trans. Pattern Analysis Machine Intelligence, 11(12): 1310–1312.

Longuet-Higgins, H.C., 1981. A computer algorithm for reconstructing scene from two projections. Nature, 293: 133–135.

Luong, Q.-T. and Faugeras, O.D., 1994. A stability analysis of the fundamental matrix. In: J.-O. Eklundh (Editor), Computer Vision-ECCV'94, Lecture Notes in Computer Science, No. 800. Springer Verlag, pp. 577–588.

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

Slama, C.C., 1980. Manual of Photogrammetry, Fourth Edition. American Society of Photogrammetry.

Thompson, E.H., 1968. The projective theory of relative orientation. Photogrammetria, 23: 67–75.

Wang, Y.N., 1995. Ein neues Verfahren zur automatischen gegenseitigen Orientierung der digitalen Bilder. Z. Photogramm. Fernerkundung, (3): 122–130.

Reprinted from

PHOTOGRAMMETRY & REMOTE SENSING

ISPRS Journal of Photogrammetry & Remote Sensing 51 (1996) 299–307

An algorithm for object reconstruction without interior orientation

J. Shan

Department of Geodetic Science, Stuttgart University, Keplerstr. 11, D-70174 Stuttgart, Germany

ELSEVIER