ISPRS Journal of Photogrammetry and Remote Sensing 66 (2011) 809–817

Contents lists available at SciVerse ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Direct relative orientation with four independent constraints

Yongjun Zhang *, Xu Huang, Xiangyun Hu, Fangqi Wan, Liwen Lin

School of Remote Sensing and Information Engineering, Wuhan University, No. 129 Luoyu Road, Wuhan 430079, PR China

ARTICLE INFO

Article history:
Received 30 April 2010
Received in revised form 7 September 2011
Accepted 7 September 2011

Keywords:
Direct relative orientation
Essential matrix
Constraint
Accuracy analysis
Least squares adjustment

ABSTRACT

Relative orientation based on the coplanarity condition is one of the most important procedures in photogrammetry and computer vision. The conventional relative orientation model has five independent parameters if interior orientation parameters are known. The model of direct relative orientation contains nine unknowns to establish the linear transformation geometry, so there must be four independent constraints among the nine unknowns. To eliminate the influence of over parameterization of the conventional direct relative orientation model, a new relative orientation model with four independent constraints is proposed in this paper. The constraints are derived from the inherent orthogonal property of the rotation matrix of the right image of a stereo pair. These constraints are completely new as compared with the known literature. The proposed approach can find the optimal solution under least squares criteria. Experimental results show that the proposed approach is superior to the conventional model of direct relative orientation, especially at low altitude and close range photogrammetric applications.

1. Introduction

Relative orientation is the process of recovering the position and orientation of one image with respect to the other in a certain image space coordinate system (Läbe and Förstner, 2006). It is a classic topic in both photogrammetry and computer vision communities (Huang and Faugeras, 1989; Faugeras and Maybank, 1990; Wang, 1990; Philip, 1996; Zhang, 1998; Mikhail et al., 2001; Stewénius et al., 2006). A pioneer eight point algorithm of relative orientation in computer vision, which is quite similar to the conventional model of direct relative orientation in photogrammetry, is proposed by Higgins (1981) although constraints among the eight unknowns are not considered. Many attentions are concentrated on resolving the fundamental or essential matrix with five conjugate points (Huang and Faugeras, 1989; Faugeras and Maybank, 1990; Philip, 1996; Nistér, 2004; Stewénius et al., 2006). Faugeras and Maybank (1990) proved that the five point algorithm has at most 10 solutions. Because of noises in the image coordinates, the essential matrix will not be exactly decomposable. This may introduce large errors in the estimation of rotation and translation parameters. Better accuracy can be achieved if the decomposability constraints are imposed (Huang and Faugeras, 1989). An iterative method was used by Horn (1990) to determine baseline and rotation parameters with an initial guess of rotation angles. However, initial guesses of rotation angles are not always reasonable in all cases especially at low altitude and close range applications. Nistér (2004) and Stewénius et al. (2006) improved the five point algorithm so that it can operate correctly even in the case of critical surfaces. The five point algorithm proposed by Stewénius et al. (2006) includes six steps. Processes of establishing linear equations from the epipolar constraint, building up 10 third-order polynomial equations with the rank and trace constraint, computing the Gröbner basis on the 10×20 matrix, computing the 10×10 action matrix, parameter back-substitution with the left eigenvectors of the action matrix and five parameter decomposition are used to compute the five relative orientation parameters. Their experiments with large point sets show that the five point algorithm provides the most consistent results.

Relative orientation is also the prerequisite of bundle adjustment (Kraus, 1993; Mikhail et al., 2001; Nistér, 2004), which can achieve the best accuracy of the data (Triggs et al., 2000; Alamouri et al., 2008). However, bundle adjustment often could not obtain the globally optimal solution with inaccurate approximations of unknowns, especially when there are some outliers (Stewénius et al., 2006). There are also methods on the absolute pose determination and 3D reconstruction with point and line correspondences (Liu et al., 1990; Kumar and Hanson, 1994; Taylor and Kriegman, 1995; Li et al., 2007). In the case of photographic configurations with large oblique angles, such as low altitude and close range applications, the approximate angular elements cannot be set as zero. Therefore, direct relative orientation which needs no initial values becomes one of the best choices. However, due to the correlation characteristics among unknowns, it usually brings down the accuracy of relative orientation and even leads to erroneous solutions in some bad geometric configurations (Stewénius et al., 2006).

* Corresponding author.
E-mail address: zhangyj@whu.edu.cn (Y. Zhang).

0924-2716/$ - see front matter © 2011 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS) Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.isprsjprs.2011.09.006

Usually, homogeneous algebraic representation and singular value decomposition strategy are used in most relative orientation algorithms of computer vision, while error equation of mathematic model and iterative least squares adjustment are used in photogrammetry. In this paper, a new direct relative orientation model from the photogrammetric point of view is proposed. Different from the epipolar constraint of essential matrix in computer vision communities that taking normalized image points as observations, original focal plane coordinates of conjugate points and the corresponding focal lengths are used in the linear model. The new model is composed of four constraints together with the conventional model of direct relative orientation. The four constraints are derived from the inherent orthogonal property of rotation matrix. Overview of conventional direct relative orientation is given in the next section. Principle of ill-posed problem that caused by over parameterization is discussed in Sections 3. The proposed new model of direct relative orientation with constraints and practical solving procedures are thoroughly addressed in Sections 4 and 5, respectively. Then several experiments are performed and compared with the ground truth in detail. Finally, conclusions are briefly outlined.

2. Conventional model of direct relative orientation

Given two images of a scene taken from different viewpoints, a stereo model can be created to reestablish the original epipolar geometry. The mathematic model of relative orientation can be expressed by coplanarity equation (Wang, 1990; Mikhail et al., 2001):

$$\mathbf{F} = \begin{vmatrix} B_x & B_y & B_z \\ u & v & w \\ u' & v' & w' \end{vmatrix} = 0 \tag{1}$$

where B_x, B_y and B_z are baseline parameters of a stereo pair, $(u\ v\ w)^T = (x\ y\ -f)^T$ and $(u'\ v'\ w')^T = \mathbf{R} \cdot (x'\ y'\ -f')^T$ coordinates of conjugate points in the image space coordinate system, (x, y), (x', y') the original focal plane coordinates of conjugate points, $\mathbf{R} = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$ the rotation matrix composed of three angles φ, ω, κ, f and f' the focal lengths of two images, respectively.

Eq. (1) can be transformed into the following linear form of Eq. (2). It is similar to the basic equation of fundamental or essential matrix (Hartley and Zisserman, 2000) used in computer vision, except that f and f' explicitly exist in the following equation.

$$L_1yx' + L_2yy' - L_3yf' + L_4fx' + L_5fy' - L_6ff' + L_7xx' + L_8xy' - L_9xf' = 0 \tag{2}$$

where

$$\begin{array}{lll} L_1 = B_x \cdot c_1 - B_z \cdot a_1 & L_2 = B_x \cdot c_2 - B_z \cdot a_2 & L_3 = B_x \cdot c_3 - B_z \cdot a_3 \\ L_4 = B_x \cdot b_1 - B_y \cdot a_1 & L_5 = B_x \cdot b_2 - B_y \cdot a_2 & L_6 = B_x \cdot b_3 - B_y \cdot a_3 \\ L_7 = B_z \cdot b_1 - B_y \cdot c_1 & L_8 = B_z \cdot b_2 - B_y \cdot c_2 & L_9 = B_z \cdot b_3 - B_y \cdot c_3 \end{array} \tag{3}$$

The coefficients $L_i (i = 1, 2, \ldots, 9)$ in Eq. (2) can only be determined up to a scale, so there are eight independent parameters. Different from known methods in computer vision communities that usually set the last element of fundamental matrix as 1.0, usually L_5 of Eq. (2) is set to be 1.0 in photogrammetry for the convenience of setting up error equations and minimizing vertical parallaxes. Suppose $L_i^0 = L_i/L_5 (i = 1, 2, \ldots, 9)$ and $L_5^0 = 1$, then Eq. (2) becomes:

$$L_1^0yx' + L_2^0yy' - L_3^0yf' + L_4^0fx' + fy' - L_6^0ff' + L_7^0xx' + L_8^0xy' - L_9^0xf' = 0 \tag{4}$$

It is the basic mathematic model of conventional direct relative orientation. This model can be used to directly calculate the eight unknowns $L_1^0, L_2^0, L_3^0, L_4^0, L_6^0, L_7^0, L_8^0, L_9^0$ without initial values. The baseline parameter B_x has no influence on building up a stereo model, it can be assumed to be known, for example the mean x-parallax. So parameters, B_y and B_z can be obtained by the following equations:

$$\begin{aligned} & L_5^2 = 2B_x^2/(L_1^{0^2} + L_2^{0^2} + L_3^{0^2} + L_4^{0^2} + L_5^{0^2} + L_6^{0^2} - L_7^{0^2} - L_8^{0^2} - L_9^{0^2}) \\ & L_i = L_i^0 \cdot L_5 (i = 1, 2, \cdots, 9) \\ & B_y = -(L_1L_7 + L_2L_8 + L_3L_9)/B_x \\ & B_z = (L_4L_7 + L_5L_8 + L_6L_9)/B_x \end{aligned} \tag{5}$$

Elements of the rotation matrix $\mathbf{R}$ can be computed by Eqs. (3) and (5). Finally, three rotation angles can be decomposed by the definition of $\mathbf{R}$ (Wang, 1990; Mikhail et al., 2001). There are two sets of possible solutions about the rotation angles ϕ, ω, κ. One of the two solutions is the true configuration, another one is the twisted pair by rotating the right image 180 degrees around the baseline. It is very easy to find the correct solution with the fact that the photographic object is in front of the camera.

3. Ill-Posed problem caused by over parameterization

Over parameterization usually results in ill-posed problem when constraint relationships among unknowns are not considered (Faugeras and Maybank, 1990). Small errors in observations may be enlarged and therefore the solution is often seriously biased from the ground truth. Suppose there is a mathematic model with the following form:

$$F_1(\mathbf{X_F}) = 0 \tag{6}$$

where $\mathbf{X_F}$ is a n-dimensional vector, i.e., there are n independent parameters in the above model. A new m-dimensional $m \geqslant n$ vector $\mathbf{Y_F}$ can be obtained by applying a certain transformation $\mathbf{Y_F} = T(\mathbf{X_F})$. So we can get the following model that takes $\mathbf{Y_F}$ as parameter:

$$F_2(\mathbf{Y_F}) = 0 \tag{7}$$

The model $F_2(\mathbf{Y}_F = 0)$ is over parameterized in the case of $m > n$. There must be $m - n$ conditional equations $G(\mathbf{Y_F}) = 0$ among the elements of vector $\mathbf{Y_F}$. It is a typical model of adjustment with functional constraints when solving the equations $F_2(\mathbf{Y_F}) = 0$ and the conditional equations $G(\mathbf{Y_F}) = 0$. The unbiased least squares solution is:

$$\mathbf{Y_1} = \mathbf{N}_{\mathbf{B}B}^{+}\mathbf{B_F^T l_F} - \mathbf{N}_{\mathbf{B}B}^{+}\mathbf{C_F^T N_{CC}^{-1} C_F N}_{\mathbf{B}B}^{+}\mathbf{B_F^T l_F} - \mathbf{N}_{\mathbf{B}B}^{+}\mathbf{C_F^T N_{CC}^{-1} W_F} \tag{8}$$

where $\mathbf{Y_1}$ is the solution of the unknown vector $\mathbf{Y_F}$, $\mathbf{N}_{\mathbf{B}B}^{+}$ is the pseudo-inverse (Hartley and Zisserman, 2000) of $(\mathbf{B_F^T B_F})$, i.e., $\mathbf{N}_{\mathbf{B}B}^{+} = (\mathbf{B_F^T B_F})^{+}$, $\mathbf{N}_{CC}^{-1} = (\mathbf{C_F N_{BB}^{+} C_F^T})^{-1}$, $\mathbf{B_F}$ is the design matrix of the unknown vector $\mathbf{Y_F}$, $\mathbf{C_F}$ is the coefficient matrix of the unknown vector $\mathbf{Y_F}$ corresponded to the conditional equations $G(\mathbf{Y_F}) = 0$, $\mathbf{l_F}$ and $\mathbf{W_F}$ are the vectors of constant items of the two kinds of equations which can be calculated by the observations and the approximate values of unknowns. The approximate value of $\mathbf{Y_F}$ has to be provided. Usually it can be set to be zero in linear cases.

However, the adjustment model to solve $F_2(\mathbf{Y_F}) = 0$ is a typical model of adjustment of observation equations if the condition $G(\mathbf{Y_F}) = 0$ is not considered. The least squares solution is just the first item of the right side of Eq. (8), i.e., $\mathbf{Y_2} = \mathbf{N}_{\mathbf{B}B}^{+}\mathbf{B_F^T l_F}$. In the case of over parameterization, especially when there are also some outliers in observations, the least squares solution can be a biased estimation if the constraints among the unknown parameters are not been considered. The conventional model of direct relative orientation is obviously over parameterized, which means the achieved solution is not always reliable.

4. Model of direct relative orientation with constraints

As described in Section 2, there are eight unknowns $L_1^0, L_2^0, L_3^0, L_4^0, L_6^0, L_7^0, L_8^0, L_9^0$ in the conventional model of direct relative orientation. As can be seen in Eq. (5), the scaling degeneracy of the nine coefficients does not exist anymore as long as B_x is fixed. In this section, we will start form Eq. (2) that contains nine unknowns to derive the new model. As we know, there should be only five independent parameters (B_y, B_z, φ, ω, κ) in the model of relative orientation if the interior parameters are known (Tang and Heipke, 1996; Zhang, 1998). So there must be four independent constraints among the nine unknowns.

Usually the well known epipolar equation $\mathbf{x}'^{\mathbf{T}}\mathbf{E}x = 0$, rank-two constraint det($\mathbf{E}$) = 0, and trace constraint $2\mathbf{E}E^{\mathrm{T}}\mathbf{E}-$ trace$(\mathbf{E}E^{\mathrm{T}})\mathbf{E} = 0$ are used in computer vision to solve the linear model of essential matrix $\mathbf{E}$ (Nistér, 2004; Stewénius et al., 2006). Here $\mathbf{x}'$ and $\mathbf{x}$ are the normalized image coordinates of conjugate points, $\mathbf{E} = \mathbf{TR}$ with $\mathbf{T}$ and $\mathbf{R}$ the translation and rotation matrix of relative orientation. Actually, the essential matrix $\mathbf{E}$ has close relation with the $L_i(i = 1, 2, \ldots, 9)$ coefficients except that the subscripts need to be switched, since both of them are based on the coplanarity constraint.

The above trace constraint is deduced from the orthogonal property of rotation matrix $\mathbf{R}R^{\mathbf{T}} = \mathbf{I}$ but with matrix form. As can be seen, the components of the trace constraint yield nine homogeneous polynomial equations of degree 3 which must be satisfied by the coefficients of $\mathbf{E}$ (Faugeras and Maybank, 1990). However, it is well known that there are five independent unknowns in relative orientation and thus the 10 constraints should be theoretically correlated. Actually, the rank-two constraint and the trace constraint used by Stewénius et al. (2006) are correlated with each other, which have already been proved by Faugeras and Maybank (1990). Detailed proof of correlation among the nine homogeneous polynomial equations of the trace constraint will be given in the following.

Suppose the essential matrix $\mathbf{E}$ has the following general form:

$$\mathbf{E} = \begin{pmatrix} e_1 & e_2 & e_3 \\ e_4 & e_5 & e_6 \\ e_7 & e_8 & e_9 \end{pmatrix} = \begin{bmatrix} 0 & -B_z & B_y \\ B_z & 0 & -B_x \\ -B_y & B_x & 0 \end{bmatrix} \mathbf{R} = \mathbf{T}R \tag{9}$$

where B_x, B_y and B_z are the translation parameters, $\mathbf{T}$ the translation matrix of relative orientation, $\mathbf{R}$ the rotation matrix of the right image of a stereo pair.

The product $\mathbf{E}E^{\mathrm{T}}$ can be obtained with the following form:

$$\begin{aligned} \mathbf{E}E^{\mathrm{T}} &= \begin{bmatrix} B_z^2 + B_y^2 & -B_xB_y & -B_xB_z \\ -B_xB_y & B_x^2 + B_z^2 & -B_yB_z \\ -B_xB_z & -B_yB_z & B_x^2 + B_y^2 \end{bmatrix} \\ &= \begin{pmatrix} e_1^2 + e_2^2 + e_3^2 & e_1e_4 + e_2e_5 + e_3e_6 & e_1e_7 + e_2e_8 + e_3e_9 \\ e_1e_4 + e_2e_5 + e_3e_6 & e_4^2 + e_5^2 + e_6^2 & e_4e_7 + e_5e_8 + e_6e_9 \\ e_1e_7 + e_2e_8 + e_3e_9 & e_4e_7 + e_5e_8 + e_6e_9 & e_7^2 + e_8^2 + e_9^2 \end{pmatrix} \end{aligned} \tag{10}$$

The trace constraint $2\mathbf{E}E^{\mathrm{T}}\mathbf{E}$ – trace$(\mathbf{E}E^{\mathrm{T}})\mathbf{E} = 0$ has the following form:

$$\begin{aligned} &2\begin{bmatrix} B_z^2 + B_y^2 & -B_xB_y & -B_xB_z \\ -B_xB_y & B_x^2 + B_z^2 & -B_yB_z \\ -B_xB_z & -B_yB_z & B_x^2 + B_y^2 \end{bmatrix} \begin{pmatrix} e_1 & e_2 & e_3 \\ e_4 & e_5 & e_6 \\ e_7 & e_8 & e_9 \end{pmatrix} \\ &\quad - 2\left(B_x^2 + B_y^2 + B_z^2\right) \begin{pmatrix} e_1 & e_2 & e_3 \\ e_4 & e_5 & e_6 \\ e_7 & e_8 & e_9 \end{pmatrix} = 0 \end{aligned} \tag{11}$$

The following nine constraints can be deduced when the above equation is expanded.

$$\begin{aligned} &(B_z^2 + B_y^2)e_1 - B_xB_ye_4 - B_xB_ze_7 = (B_x^2 + B_y^2 + B_z^2)e_1 \\ &(B_z^2 + B_y^2)e_2 - B_xB_ye_5 - B_xB_ze_8 = (B_x^2 + B_y^2 + B_z^2)e_2 \\ &(B_z^2 + B_y^2)e_3 - B_xB_ye_6 - B_xB_ze_9 = (B_x^2 + B_y^2 + B_z^2)e_3 \\ &-B_xB_ye_1 + (B_x^2 + B_z^2)e_4 - B_yB_ze_7 = (B_x^2 + B_y^2 + B_z^2)e_4 \\ &-B_xB_ye_2 + (B_x^2 + B_z^2)e_5 - B_yB_ze_8 = (B_x^2 + B_y^2 + B_z^2)e_5 \\ &-B_xB_ye_3 + (B_x^2 + B_z^2)e_6 - B_yB_ze_9 = (B_x^2 + B_y^2 + B_z^2)e_6 \\ &-B_xB_ze_1 - B_yB_ze_4 + (B_x^2 + B_y^2)e_7 = (B_x^2 + B_y^2 + B_z^2)e_7 \\ &-B_xB_ze_2 - B_yB_ze_5 + (B_x^2 + B_y^2)e_8 = (B_x^2 + B_y^2 + B_z^2)e_8 \\ &-B_xB_ze_3 - B_yB_ze_6 + (B_x^2 + B_y^2)e_9 = (B_x^2 + B_y^2 + B_z^2)e_9 \end{aligned} \tag{12}$$

It can be easily deduced that there are only three independent ones among the above nine constraints when correlated equations are removed.

$$\begin{aligned} -B_ye_4 - B_ze_7 &= B_xe_1 \\ -B_ye_5 - B_ze_8 &= B_xe_2 \\ -B_ye_6 - B_ze_9 &= B_xe_3 \end{aligned} \tag{13}$$

Multiplying B_x on both sides of Eq. (13), replacing $-B_xB_y$ and $-B_xB_z$ by the corresponding items in Eq. (10), we can get the following three constraints:

$$\begin{aligned} (e_1e_4 + e_2e_5 + e_3e_6)e_4 + (e_1e_7 + e_2e_8 + e_3e_9)e_7 &= B_x^2e_1 \\ (e_1e_4 + e_2e_5 + e_3e_6)e_5 + (e_1e_7 + e_2e_8 + e_3e_9)e_8 &= B_x^2e_2 \\ (e_1e_4 + e_2e_5 + e_3e_6)e_6 + (e_1e_7 + e_2e_8 + e_3e_9)e_9 &= B_x^2e_3 \end{aligned} \tag{14}$$

Up to now, it is clear that the nine homogeneous polynomial equations of the trace constraint are correlated.

Different from the epipolar constraint of essential matrix in computer vision communities that taking normalized image points as observations, original focal plane coordinates of conjugate points are used in this paper. Focal lengths f and f' are also remained in Eq. (2). The four independent constraints of the proposed method will be directly deduced from the orthogonal property of rotation matrix $\mathbf{R}$ in the following.

The orthogonal property of $\mathbf{R}$ can be ensured by the following equations:

$$\begin{aligned} a_1^2 + a_2^2 + a_3^2 = 1 \quad & a_1b_1 + a_2b_2 + a_3b_3 = 0 \\ b_1^2 + b_2^2 + b_3^2 = 1 \quad & a_1c_1 + a_2c_2 + a_3c_3 = 0 \\ c_1^2 + c_2^2 + c_3^2 = 1 \quad & b_1c_1 + b_2c_2 + b_3c_3 = 0 \end{aligned} \tag{15}$$

where $a_i, b_i, c_i (i = 1, 2, 3)$ are the nine elements of the rotation matrix $\mathbf{R}$ which is already described in Eq. (1). The following four formulae can be obtained from the expression of $L_i(i = 1, 2, \ldots, 9)$ in Eq. (3) and the orthogonal property of rotation matrix in Eq. (15):

$$\begin{aligned} &L_1^2 + L_2^2 + L_3^2 = B_x^2 + B_z^2 \\ &L_4^2 + L_5^2 + L_6^2 = B_x^2 + B_y^2 \\ &L_7^2 + L_8^2 + L_9^2 = B_y^2 + B_z^2 \\ &L_1L_4 + L_2L_5 + L_3L_6 = B_y \cdot B_z \end{aligned} \tag{16}$$

As described in Section 2, the baseline parameter B_x can be defined as the mean x-parallax of conjugate points while calculating the initial values of nine coefficients $L_i(i = 1, 2, \ldots, 9)$. Taking the expressions of B_y and B_z in Eq. (5) into account, the four independent constraints among the nine coefficients of Eq. (2) can be obtained:

$$
\begin{aligned}
&L_1^2+L_2^2+L_3^2=B_x^2+(L_4L_7+L_5L_8+L_6L_9)^2/B_x^2\\
&L_4^2+L_5^2+L_6^2=B_x^2+(L_1L_7+L_2L_8+L_3L_9)^2/B_x^2\\
&L_7^2+L_8^2+L_9^2=(L_1L_7+L_2L_8+L_3L_9)^2/B_x^2+(L_4L_7+L_5L_8+L_6L_9)^2/B_x^2\\
&L_1L_4+L_2L_5+L_3L_6=-(L_1L_7+L_2L_8+L_3L_9)\cdot(L_4L_7+L_5L_8+L_6L_9)/B_x^2
\end{aligned}
\tag{17}
$$

The above four equations are the constraints among the nine unknowns of direct relative orientation. So the proposed new model of direct relative orientation with constraints is composed of Eqs. (2) and (17). The proposed new model chooses the original nine parameters of direct relative orientation as unknowns, and four constraints are combined to avoid the problem of over parameterization.

5. Practical procedure of solving the problem

General procedure of solving relative orientation parameters of the proposed new model is composed of three steps. Firstly, the conventional direct relative orientation is performed to get approximate values of the nine unknowns. Then four constraints are combined into the conventional model to avoid the problem of over parameterization. Finally, the five relative orientation parameters are decomposed by the computed nine unknowns.

Dividing both sides of Eq. (2) by L_5f, adding y to both sides and moving y' to the right side, one can obtain the nonlinear equation of vertical parallax:

$$
\frac{yx'}{f}\frac{L_1}{L_5}+\frac{yy'}{f}\frac{L_2}{L_5}-\frac{yf'}{f}\frac{L_3}{L_5}+\frac{fx'}{f}\frac{L_4}{L_5}+y-\frac{ff'}{f}\frac{L_6}{L_5}+\frac{xx'}{f}\frac{L_7}{L_5}+\frac{xy'}{f}\frac{L_8}{L_5}-\frac{xf'}{f}\frac{L_9}{L_5}=y-y' \tag{18}
$$

Note that the above equation minimizes vertical parallaxes between images. In the cases of geometric configurations with vertical epipolar lines, the above equation fails to solve the problem. However, this can be resolved by minimizing the horizontal parallaxes. $L_4^0=1$ can be assumed in the conventional model to obtain the initial values of the nine coefficients with the similar strategy to that in Section 2. Then dividing both sides of Eq. (2) by L_4f, adding x to both sides and moving x' to the right side, one can obtain the nonlinear equation of horizontal parallax:

$$
\frac{yx'}{f}\frac{L_1}{L_4}+\frac{yy'}{f}\frac{L_2}{L_4}-\frac{yf'}{f}\frac{L_3}{L_4}+x+\frac{fy'}{f}\frac{L_5}{L_4}-\frac{ff'}{f}\frac{L_6}{L_4}+\frac{xx'}{f}\frac{L_7}{L_4}+\frac{xy'}{f}\frac{L_8}{L_4}-\frac{xf'}{f}\frac{L_9}{L_4}=x-x' \tag{19}
$$

Each pair of conjugate image points provides one observation equation. Therefore at least nine pairs of conjugate points are required to solve the above equations. The general error equation of Eqs. (18) and (19) can be linearized into the following form according to the model of adjustment of observation equations:

$$
\mathbf{V}=\mathbf{B}X-\mathbf{l} \tag{20}
$$

where $\mathbf{V}$ is the correction vector of all observations, $\mathbf{B}$ the design matrix of unknown vector $\mathbf{X}$ and $\mathbf{1}$ the vector of constant items of error equations which can be calculated by observations and approximate values of unknowns.

The constraint Eq. (17) can be linearized as:

$$
\mathbf{C}X-\mathbf{W}_X=\mathbf{0} \tag{21}
$$

The elements of coefficient matrix $\mathbf{C}$ can be obtained by the partial derivatives of corresponding unknowns, $\mathbf{W}_X$ is the vector of constant items by incorporating the approximate values of $L_i(i=1,2,\ldots,9)$ into Eq. (17).

The mathematical model of the proposed method is the combination of Eqs. (20) and (21). According to the principle of least squares adjustment with functional constraints (Mikhail et al., 2001), the normal equation can be written as follows when all observations are assumed to have the same weight (Wang, 1990; Mikhail et al., 2001):

$$
\begin{bmatrix}\mathbf{B}^{\mathrm{T}}\mathbf{B} & \mathbf{C}^{\mathrm{T}}\\ \mathbf{C} & \mathbf{f}0\end{bmatrix}\cdot\begin{bmatrix}\mathbf{X}\\ \mathbf{K}\end{bmatrix}-\begin{bmatrix}\mathbf{B}^{\mathrm{T}}\mathbf{l}\\ \mathbf{W}_X\end{bmatrix}=\mathbf{0} \tag{22}
$$

The unknown vector $\mathbf{X}$ and connection vector $\mathbf{K}$ can be resolved by inversing the above normal equation. We can get the correction vector $\mathbf{V}$ by substituting $\mathbf{X}$ into error Eq. (20). Then the precision evaluation and gross error detection can be performed by analyzing the residues of observations and the variance and co-variance matrix which can be computed by the inversion of the normal matrix of Eq. (22). Note that the computing of $L_i(i=1,2,\ldots,9)$ coefficients is an iterative process to detect outliers and get precise solution. Iteration is terminated if the root mean square error of unit weight or the corrections of all unknowns are smaller than the given threshold. Finally, the five elements $(B_y, B_z, \varphi, \omega, \kappa)$ of relative orientation can be decomposed with the same strategy as that of the conventional relative orientation (Wang, 1990; Mikhail et al., 2001).

In summary, our approach includes three steps: getting an initial solution of the conventional linear model; improvement of the nine coefficients by incorporating the derived four constraints; and decomposition of the five parameters with the nine coefficients. As compared with the five point algorithm proposed by Stewénius

Fig. 1. Overview of the photographic area of aerial images.

et al. (2006) which includes six steps and manipulations of 10×20 matrices, the computation efficiency of our approach is also superior.

6. Experiments and results

To verify the correctness and effectiveness of the proposed new model of direct relative orientation with constraints, three test data sets of conjugate points from aerial, low altitude and terrestrial close range images are used for experiments, respectively. For each type of images, conventional direct relative orientation and the proposed direct relative orientation with constraints are both performed and compared with the ground truth obtained by bundle adjustment.

6.1. Results of relative orientation with aerial images

Aerial images of 20 stereo pairs are taken by Z/I Imaging DMC camera with pixel size of 12 μm. The photographic area is shown in Fig. 1. Mean ground resolution of the aerial images is about 0.25 m. In order to get reliable relative orientation results with redundant observations, more than 100 conjugate points are

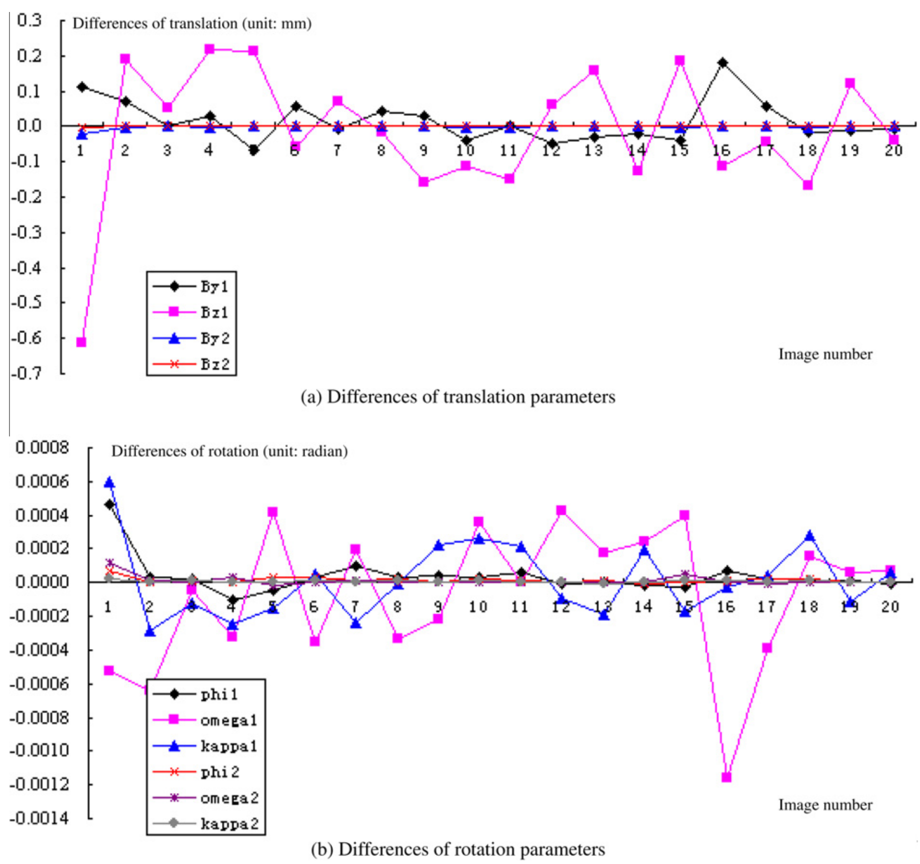

Fig. 2. Differences between results of the two methods and the ground truth with aerial images. 1 in the above figure means results of the conventional method while 2 means those of the new method.

Fig. 3. Overview of the photographic area of low altitude images.

automatically matched for each stereo pair, although only nine is theoretically necessary to calculate the nine coefficients. There is no outlier in conjugate points since the test data has already been processed by bundle adjustment. Unit weight root mean square errors of each stereo pair for the two methods mentioned above are all smaller than 0.3 pixels. The ground truth of relative orientation parameters are calculated from the results of bundle adjustment.

For the convenience of comparison, baseline length B_x is fixed as the mean x-parallax of conjugate points of each stereo pair for both the two methods. It is around 36.0 mm since the nominal overlap between adjacent images is 60%. Differences between results of the two methods and the ground truth are given in Fig. 2. By1 and Bz1 represent the differences of baseline parameters B_y and B_z of the conventional method against the ground truth; phi1, omega1 and kappa1 represent the differences of angular parameters φ, ω, and κ of the conventional method against the ground truth. By2, Bz2, phi2, omega2 and kappa2 represent the differences of corresponding translation and rotation parameters of the new method against the ground truth. Units of translation and rotation parameters are mm and radian, respectively.

It can be seen from Fig. 2 that the differences between results of the proposed method and the ground truth are much smaller than that between results of the conventional method and the ground truth. It shows that although the conventional direct relative

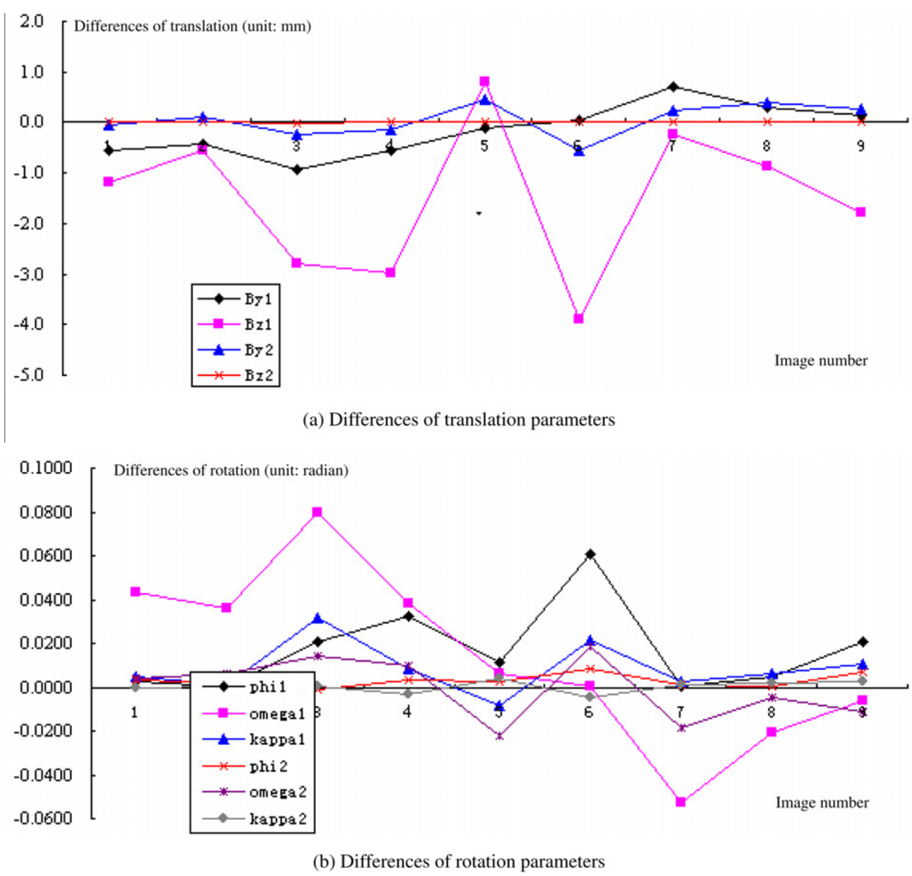

Fig. 4. Differences between results of the two methods and the ground truth with low altitude images. 1 in the above figure means results of the conventional method while 2 means those of the new method.

Fig. 5. Overview of the terrestrial photographic object.

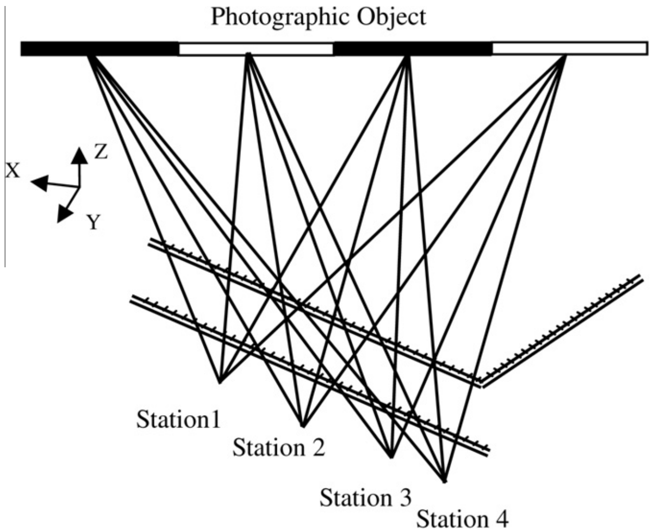

Fig. 6. Photographic stations of terrestrial object and convergent angles.

orientation requires no initial values of unknown parameters, the precision of decomposed five elements cannot be guaranteed in all cases. Sometimes it will generate imprecise relative orientation parameters because of the over parameterization problem, such as the results of the first and 16th stereo pair. However, the proposed new model starts from the results of conventional direct relative orientation and takes additional constraints into account while doing least squares adjustment to overcome the problem of over parameterization, so the computed relative orientation parameters are closer to the ground truth.

6.2. Results of relative orientation with low altitude images

Low altitude images of nine stereo pairs are acquired by a pre-calibrated non-metric digital camera Kodak Pro SLR with 24×36 mm image format and 8 μm pixel sizes. The low altitude images are photographed with 80% forward overlap by an unmanned airship on which the camera with 24 mm lens is mounted. Flying height is 150 m above the ground, so the ground resolution is about 0.05 m. The photographic area of low altitude images is shown in Fig. 3. There are significant projection distortions of photographed objects above the ground, such as buildings and trees. Conjugate points are obtained by automatic image matching without any further processing, such as relative orientation and bundle adjustment, about 2% of them are outliers. The minimum number of conjugate points of a stereo pair is 126. The two methods mentioned above are used for experiments with these low altitude images. Robust estimation with data snooping technique (Baarda, 1968) is adopted to detect and remove outliers for both the two methods. The ground truth of relative orientation parameters is calculated from the results of bundle adjustment. B_x of each model is also given by mean x-parallax. It is around 5.0 mm since the overlap of the adjacent images is about 80%. There are significant changes of overlap and orientation angles between adjacent images. Unit weight root mean square errors of all stereo pairs for the two methods are all smaller than 0.5 pixels.

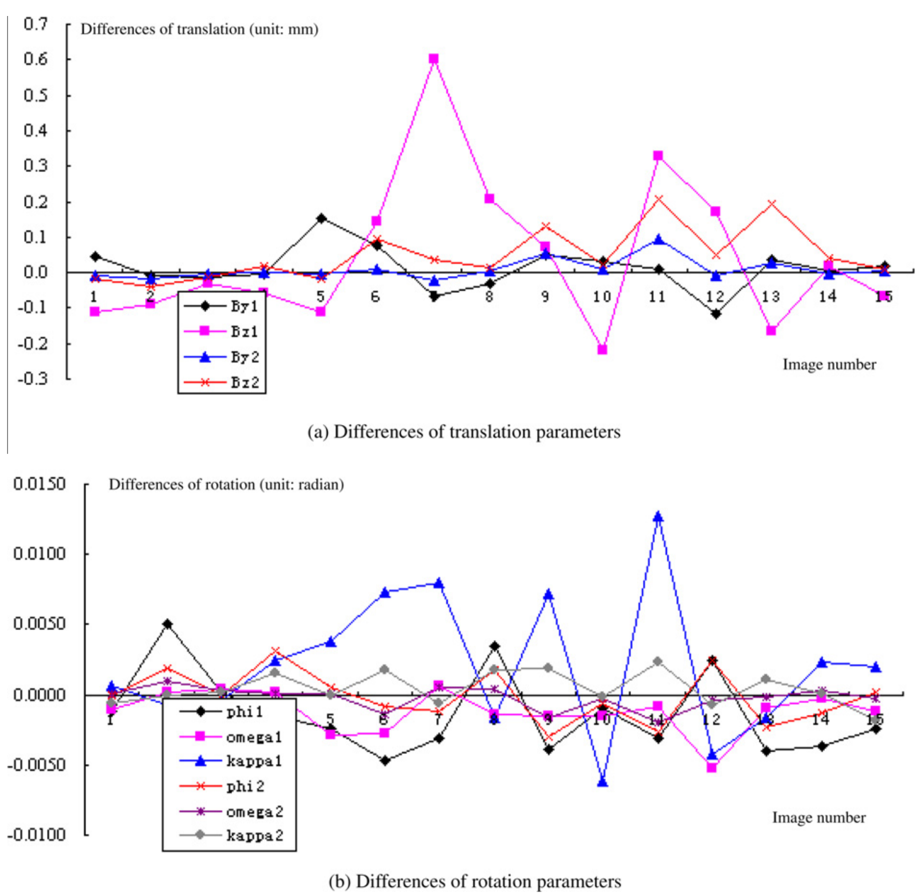

Fig. 7. Differences between results of the two methods and the ground truth with terrestrial close range images. 1 in the above figure means results of the conventional method while 2 means those of the new method.

Differences between results of the two methods and the ground truth are given in Fig. 4. Definition of each symbol in Fig. 4 is the same as that in Fig. 2. It can be seen that the differences of baseline component B_z of the stereo pairs 3, 4 and 6 obtained by the conventional method are larger than a half of B_x, which is only about 5.0 mm as mentioned above. Obviously, the computed results of the aforementioned three models are completely wrong. It shows that the conventional direct relative orientation cannot get reasonable relative orientation parameters when there are large variations of relative orientation parameters together with outliers. However, direct relative orientation with constraints can give more stable and reasonable results of relative orientation. Results of the proposed method are precise enough to be used as initial values for subsequent computations, such as bundle block adjustment.

6.3. Results of relative orientation with terrestrial close range images

Terrestrial close range images of 15 stereo pairs are acquired by the same camera as that in Section 6.2 but with 50 mm lens. The terrestrial photographic object, an engineering scene, is shown in Fig. 5. There are totally four photographic stations. Four images are taken from each station, so totally 16 images are photographed. Fig. 6 shows the photographic stations and directions of principle axes of each station. The mean photographic distance is about 230 m, so the mean ground resolution is 0.036 m. The minimum number of conjugate points of a stereo pair is 206. There are no outliers in conjugate points since the test data has already been processed by bundle adjustment with enough ground control points. The two methods mentioned above are used for experiments with these close range images. B_x of each model is also computed by mean x-parallax. There are significant changes of translation and rotation angles between adjacent images because of large convergent angle. The ground truth of relative orientation parameters is also calculated from the results of bundle adjustment. The unit weight root mean square error of each model is around 0.5 pixels.

Differences between results of the two methods and the ground truth are given in Fig. 7. Definition of each symbol in Fig. 7 is the same as that in Fig. 2. As can be seen, the maximum difference of relative orientation parameters between the conventional method and the ground truth is about 0.6 mm and 0.013 rad, respectively, while the maximum difference of relative orientation parameters between the proposed method and the ground truth is only 0.2 mm and 0.007 rad, respectively. It shows that the result of direct relative orientation with constraints is also quite better than that of the conventional direct relative orientation.

7. Conclusions

A new approach of direct relative orientation with constraints based on the conventional model of direct relative orientation is proposed in this paper. The proposed new approach incorporates four non-linear constraints among the nine unknowns of the conventional model of direct relative orientation to reduce the over parameterization problem. The constraints are derived from the inherent orthogonal property of rotation matrix. General procedure of the proposed method includes three steps, conventional direct relative orientation, least squares adjustment with constraints, and finally five relative orientation parameters decomposition. Gross error detection with data snooping techniques is applied by analyzing the residues of observations and the variance and co-variance matrix of unknowns, which can be computed by inversion of the normal matrix.

In terms of aerial and terrestrial stereos with no outliers, results of the conventional method and the proposed method are both reasonable when compared with the ground truth. However, the results of the proposed method are quite better than that of the conventional method in some cases. In the case of conjugate points of low altitude images with a few percent outliers, some results of the conventional method are completely wrong. However, the proposed method can give reliable results without any human interaction. That means the proposed method can be used in applications where no initial information available, such as close range or low altitude oblique photography.

Ten constraints are usually used on the nine elements of essential matrix in computer vision literatures. However, it is well known that there are only five independent unknowns in relative orientation if the intrinsic parameters are known, so there should be definitely four independent constraints among the nine elements. The proof, that the 10 constraints in the known literature are related and only four of them are independent, are given for the first time. It seems that the four but not the 10 constraints should be used to recover the essential matrix.

Generally, the limitation of our method is similar to that of the conventional eight-point algorithms. It is possible to fail in degenerate cases of computer vision applications. However, degenerate situations such as planar scenes and critical surfaces are not common for photogrammetry since it mainly deals with natural terrains. In the general over determined cases, the performance of our method is theoretically superior to the five point algorithm and the conventional eight-point algorithm because the correct number of constraints is applied. Further work will be made to fully investigate the differences, performances, advantages and disadvantages about the proposed model and the available five point algorithm.

Acknowledgements

This work is supported by National Natural Science Foundation of China with Project Nos. 41071233 and 41171292, National Key Technology Research and Development Program with Project No. 2011BAH12B05, National Hi-Tech Research and Development Program with Project No. 2009AA121403, Non-profit Special Project of Ministry of Land and Resources with Project No. 201011020-4. Heartfelt thanks are also given for the comments and contributions of anonymous reviewers especially on the property of essential matrix.

References

Alamouri, A., Gruendig, L., Kolbe, T.H., 2008. A new approach for relative orientation of non-calibrated historical photos of Baalbek/Libanon. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 37 (Pt. B5), 279–284.

Baarda, W., 1968. A testing procedure for use in geodetic networks. Publications on Geodesy 2(5), 1–97, Delft.

Faugeras, O., Maybank, S., 1990. Motion from point matches: multiplicity of solutions. International Journal of Computer Vision 4 (3), 225–246.

Hartley, R. I., Zisserman, A., 2000. Multiple View Geometry in Computer Vision. Cambridge University Press.

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

Horn, B.K.P., 1990. Relative orientation. International Journal of Computer Vision 4 (1), 59–78.

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

Kumar, R., Hanson, A.R., 1994. Robust methods for estimating pose and a sensitivity analysis. CVGIP: Image Understanding 60 (3), 313–342.

Kraus, K., 1993. Photogrammetry. Volume 1, Fundamentals and Standard Processes. Dummler Verlag, Bonn.

Läbe, T., Förstner, W., 2006. Automatic relative orientation of images. Proceedings of the 5th Turkish-German Joint Geodetic Days, 29–31 March. Berlin, Germany (on CD-ROM).

Li, Z. G., Liu, J. Z., Tang, X. O., 2007. A closed-form solution to 3D reconstruction of piecewise planar objects from single images. IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–6.

Liu, Y., Huang, T.S., Faugeras, O.D., 1990. Determination of camera location from 2-D to 3-D line and point correspondences. IEEE Transactions on Pattern Analysis and Machine Intelligence 12 (1), 28–37.
Mikhail, E.M., Bethel, J.S., McGlone, J.C., 2001. Introduction to Modern Photogrammetry. John Wiley & Sons, Inc.
Nistér, D., 2004. An efficient solution to the five-point relative pose problem. IEEE Transactions on Pattern Analysis and Machine Intelligence 26 (6), 756–770.
Philip, J., 1996. A non-iterative algorithm for determining all essential matrices corresponding to five point pairs. Photogrammetric Record 15 (88), 589–599.
Stewénius, H., Engels, C., Nisté, D., 2006. Recent developments on direct relative orientation. ISPRS Journal of Photogrammetry & Remote Sensing 60 (4), 284–294.
Tang, L., Heipke, C., 1996. Automatic relative orientation of aerial images. Photogrammetric Engineering & Remote Sensing 62 (1), 47–55.
Taylor, C.J., Kriegman, D.J., 1995. Structure and motion from line segments in multiple images. IEEE Transactions on Pattern Analysis and Machine Intelligence 17 (11), 1021–1032.
Triggs, B., McLauchlan, P., Hartley, R., Fitzgibbon, A., 2000. Bundle adjustment – a modern synthesis. Vision algorithms: theory and practice. Lecture Notes in Computer Science 1883, 298–372.
Wang, Z.Z., 1990. Principles of Photogrammetry With Remote Sensing. Press of Wuhan Technical University of Surveying and Mapping, Wuhan, China.
Zhang, Z.Y., 1998. Determining the epipolar geometry and its uncertainty: a review. International Journal of Computer Vision 27 (2), 161–198.