A robust linear least-squares estimation of camera exterior orientation using multiple geometric features

Qiang Ji[a] *, Mauro S. Costa[b], Robert M. Haralick[b], Linda G. Shapiro[b]

Department of Computer Science [a]
University of Nevada at Reno
Reno, NV 89667, USA
Department of Electrical Engineering[b]
University of Washington
Seattle, WA 98195, USA

Abstract

For photogrammetric applications, solutions to camera exterior orientation problem can be classified into linear (direct) and non-linear. Direct solutions are important because of their computational efficiency. Existing linear solutions suffer from lack of robustness and accuracy partially due to the fact that the majority of the methods utilize only one type of geometric entity and their frameworks do not allow simultaneous use of different types of features. Furthermore, the orthonormality constraints are weakly enforced or not enforced at all. We have developed a new analytic linear least-squares framework for determining camera exterior orientation from the simultaneous use of multiple types of geometric features. The technique utilizes 2D/3D correspondences between points, lines, and ellipse-circle pairs. The redundancy provided by different geometric features improves the robustness and accuracy of the least-squares solution. A novel way of approximately imposing orthonormality constraints on the sought rotation matrix within the linear framework is presented. Results from experimental evaluation of the new technique using both synthetic data and real images reveal its improved robustness and accuracy over existing direct methods.

Keywords: Exterior orientation; linear methods; least-squares; orthonormality constraints; feature fusion

1. Introduction

Camera exterior orientation estimation is an essential step for many photogrammetric applications. It addresses the issue of determining the exterior parameters (position and orientation) of a camera with respect to a world coordinate frame. Solutions to the exterior orientation problem can be classified into linear and non-linear methods. Linear methods have the advantage of computational efficiency, but they suffer from lack of accuracy and robustness. Non-linear methods, on the other hand, offer a more accurate and robust

*Corresponding author e-mail: qiangji@cs.unr.edu

solution. They are, however, computationally intensive and require initial estimates. The classical non-linear photogrammetric approach to exterior orientation (e.g. the bundle adjustment method) requires setting up an non-linear least-squares system. Given initial estimates of the exterior parameters, the system is then linearized and solved iteratively. While the classical technique guarantees the orthonormality of the rotation matrix and delivers the best answer, it, however, requires good initial estimates. It is a well-known fact that the initial estimates must be close or the system may not converge quickly or correctly. Hence, the quality of initial estimates is critical since it determines the convergence speed and the correctness of the iterative procedure. Robust and accurate linear solutions, which are often used to provide initial guesses for non-linear procedures, are therefore important for photogrammetric problems.

Numerous methods have been proposed to analytically obtain camera exterior parameters. Previous methods have primarily been focused on using sets of 2D-3D point correspondences to solve for the transformation matrix, followed by extracting the camera parameters from the solved transformation. The linear method using points is well-known in photogrammetry as *Direct Linear Transformation (DLT)*.

The original proposal of DLT method appears in (Abdel-Aziz and Karara, 1971). Since then, different variations of DLT methods have been introduced. For example, Bopp and Krauss (1978) published a variation of the DLT, incorporating added constraints into the solution. Okamoto (1981) gave an alternative derivation of the DLT from a more general mathematical framework. Shan (1996) introduced a linear solution for object reconstruction from a stereopair without interior orientation and less requirements on known points than the original DLT formulation.

In computer vision, the DLT-like methods include the three point solution (Fischler and Bolles, 1981), the four point solutions (Hung *et al.*, 1985; Holt and Netravali, 1991), and the six or more point solutions (Tsai, 1987; Sutherland, 1974; Faugeras, 1993). Haralick *et al.* (1994) reviewed and compared major direct solutions of exterior orientation using three point correspondences and characterized their performance under varying noisy conditions. Sutherland (1974) provided a closed-form least-squares solution using six or more points. The solution, however, is only up to a scale factor. Faugeras (1993) proposed a similar technique that solves the scale factor by applying a normality constraint. His solution also includes a post-orthogonalization process that ensures the orthonormality of the resulting rotation matrix. Tsai (1987) presented a direct solution by decoupling the camera parameters into two groups; each group is solved for separately in different stages. While efficient, Tsai's method does not impose any of the orthonormal constraints on the estimated rotation matrix. Also, the errors with the camera parameters estimated in the earlier stage can significantly affect the accuracy of parameters estimated in the later stage.

These methods are effective and simple to implement. However, they are not robust and are very susceptible to noise in image coordinates (Wang and Xu, 1996), especially when the number of control points approaches the minimum required. For the 3-point solutions, Haralick *et al.* (1994) shows that even the order of algebraic substitutions can render the output useless. Furthermore, the point distribution and noise in the point coordinates can also dramatically change the relative output errors. For least-squares based methods, a different study by Haralick *et al.* (1989) shows that when the noise level exceeds certain level or the number of points is below certain level, these methods become extremely unstable and the errors skyrocket. The use of more points can help relieve this problem. However, generation of more control points often proves to be difficult, expensive, and time-consuming. Another disadvantage of point-based methods is the difficulty with point matching, i.e., finding the correspondences between the 3D scene points and 2D image pixels.

In view of these issues, other researchers have investigated the use of higher-level geometric features such as lines or curves as observed geometric entities to improve the robustness and accuracy of linear methods for estimating exterior parameters. Over the years, various algorithms using features other than points for exterior orientation problems have been introduced both in photogrammetry and computer vision (Echigo, 1990; Liu *et al.*, 1990; Chen and Tsai, 1990; Rothwell *et al.*, 1992; Haralick and Chu, 1984; Tommaselli and Lugnani, 1988; Mulawa and Mikhail, 1988; Mulawa, 1989; Wang and Tsai, 1990; Lee *et al.*, 1990; Chen and Tsai, 1991; Heikkila, 1991; Mikhail, 1993; Finsterwalder, 1991; Doehler, 1975; Weng *et al.*, 1992; Petsa and Patias, 1994a,b; Paderes *et al.*, 1984). In photogrammetry, Strunz (1992) gives a good overview of using various features (points, lines, surfaces) for different photogrammetric tasks, including camera orientation. Szczepanski (1958) reviewed nearly 60 different solutions for space resection, dating back to 1829, for the simultaneous and separate determination of the position and rotation parameters. An iterative Kalman filtering method for space resection using straight line features is described in G.Tommaselli and Tozzi (1996). Masry (1981) described a method for absolute camera orientation and Lugnani Lugnani (1980) for camera exterior orientation by spatial resection, using linear features. Drewniok and Rohr (1997) presented an approach for automatic exterior orientation of aerial imagery which is based on detection and localisation of planar objects (manhole covers). Ethrog (1984) used parallel and perpendicular lines of objects for estimation of the rotation and interior orientation of non-metric cameras.

Researchers have also used other known geometric shapes in the scene to constrain the solution. Such shapes can be 2D (straight lines, circles etc.) or 3D (features on a plane, cylinder etc.) (Mikhail and Mulawa, 1985). Others (Kruck, 1984; Forkert, 1996; Kager, 1989) have incorporated geometric constraints such as coordinate differences, horizontal and space distances, and angles to improve the traditional bundle adjustment method. Heikkila (1990), Pettersen (1992), and Maas (1999) employed a moving reference bar (known distance) for camera orientation and calibration.

In computer vision, Haralick and Chu (1984) presented a method that solves the camera exterior parameters from the conic curves. Given the shape of conic curves, the method first solves for the three rotation parameters using an iterative procedure. The three translation parameters are then solved analytically. The advantage of this method is that it does not need to know the location of the curves and it is more robust than any analytical method in that rotation parameters errors are reduced to minimum before they are used analytically to compute translation parameters. In their analytic method, Liu *et al.* (1990) and Chen and Tsai (1990) discussed direct solutions for determining camera exterior parameters based on a set of 2D-3D line correspondences. The key to their approach lies in the linear constraint they used. This constraint uses the fact that a 3D line and its image line lie on the same plane determined by the center of perspectivity and the image line. Rothwell *et al.* (1992) discussed a direct method that determines camera parameters using a pair of conic curves. The method works by extracting 4 or 8 points from conic intersections and tangencies. Exterior camera parameters are then recovered from these points. Kumar and Hanson (1989) described a robust technique for finding camera parameters using lines. Kamgar-Parsi and Eas (1990) introduced a camera calibration method with small relative angles. Gao (1992) introduced a method for estimating exterior parameters using parallelepipeds. Forsyth *et al.* (1991) proposed to use a pair of known conics or a single known circle for determining the pose of the object plane. Haralick (1988) and Haralick and Chu (1984) presented methods for solving for camera parameters using rectangles and triangles. Abidi (1995) presented a closed form solution for pose estimation using quadrangular targets. Linnainmaa and Harwood (1988) discussed an approach for determination of three-dimensional object using triangle pairs. Chen and Tsai (1991) proposed closed solution for pose estimation from line-to-plane correspondences and studied the condition of the existence of the closed solution. Ma (1993) introduced a technique for pose estimation from the correspondence of 2D/3D conics. The technique, however, is iterative and requires a pair of conics in both 2D and 3D.

Analytic solutions based on high-level geometric features afford better stability and are more robust and accurate. Here, the correspondence problem can be addressed more easily than for the point-based methods. However, high-level geometric features may not always be present in some applications, and points are present in many applications. Therefore, completely ignoring points while solely employing high-level geometric entities can be a waste of readily available geometric information. This is one of the problems with the existing solutions: they either use points or lines or conics but not a combination of features. In this paper, we describe an integrated least-squares method that solves for the camera transformation matrix analytically by fusing available observed geometric information from different levels of abstraction. Specifically, we analytically solve for the external camera parameters from simultaneous use of 2D-3D correspondences between points, between lines, and between 2D ellipses and 3D circles. The attractiveness of our approach is that the redundancy provided by features at different levels improves the robustness and accuracy of the least-squares solution, therefore improving the precision of the estimated parameters. To our knowledge, no previous research attempts have been made in developing a linear solution to exterior orientation with simultaneous use of all three classes of features. Work by Phong *et al.* (1995) described a technique in which information from both points and lines is used to compute the exterior orientation. However, the method is iterative and involves only points and lines.

Another major factor that contributes to the lack of robustness of the existing linear methods is that orthonormality constraints on the rotation matrix are often weakly enforced or not enforced at all. In this research, we introduce a simple, yet effective, scheme for approximately imposing the orthonormal constraints on the rotation matrix. While the scheme does not guarantee that the resultant rotation matrix completely satisfies the orthonormal constraints, it does yield a matrix that is closer to orthonormality than those obtained with competing methods.

This paper is organized as follows. Section 2 briefly summarizes the perspective projection geometry and equations. Least-squares frameworks for estimating the camera trans-

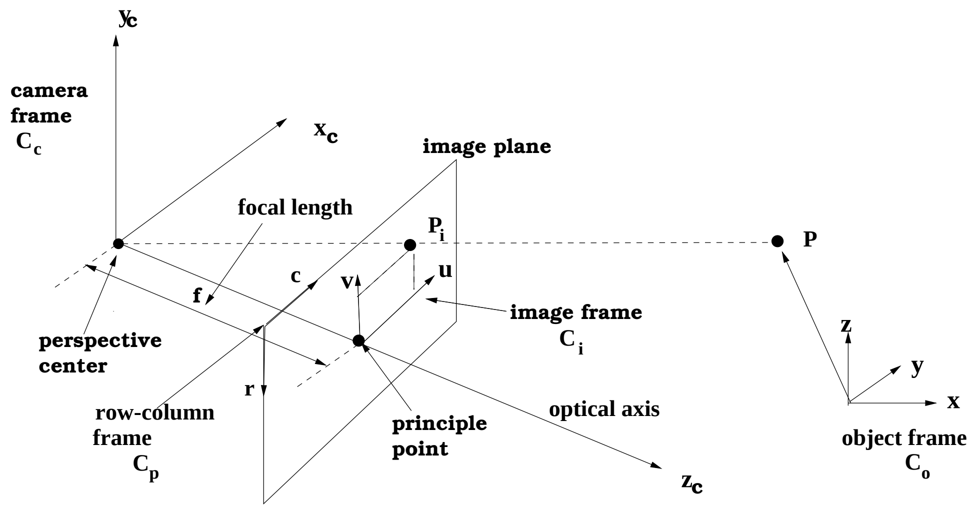

Fig. 1. Camera and perspective projection geometry.

formation matrix from 2D-3D point, line, and ellipse/circle correspondences are presented in sections 3, 4, and 5 respectively. Section 6 discusses our technique for approximately imposing orthonormal constraints and presents the integrated linear technique for estimating the transformation matrix simultaneously using point, line, and ellipse/circle correspondences. Performance characterization and comparison of the developed integrated technique is covered in section 7.

2. Perspective projection geometry

To set the stage for the subsequent discussion, this section briefly summarizes the pin-hole camera model and the perspective projection geometry.

Let P be a 3D point and $(x\ y\ z)^t$ be the coordinates of P relative to the object coordinate frame C_o. Define the camera coordinate system C_c to have its z-axis parallel to the optical axis of the camera lens and its origin located at the perspective center. Let $(x_c\ y_c\ z_c)^t$ be the coordinates of P in C_c. Define C_i to be the image coordinate system, with its u-axis and v-axis parallel to the x and y axes of the camera coordinate frame, respectively. The origin of C_i is located at the principal point. Let $(u\ v)^t$ be the coordinates of P_i, the image projection of P in C_i. Fig. 1 depicts the pin-hole camera model.

Based on the perspective projection theory, the projection that relates (u, v) on the image plane to the corresponding 3D point (x_c, y_c, z_c) in the camera frame can be described by

$$\lambda \begin{pmatrix} u \\ v \\ f \end{pmatrix} = \begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} \tag{1}$$

where λ is a scalar and f is the camera focal length.
Further, $(x\ y\ z)^t$ relates to $(x_c\ y_c\ z_c)^t$ by a rigid body coordinate transformation consisting of a rotation matrix and a translation. Let a 3×3 matrix R represent the rotation and

a 3×1 vector T describe the translation, then

$$\begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} = R \begin{pmatrix} x \\ y \\ z \end{pmatrix} + T \tag{2}$$

where T and R can be parameterized as $T = (t_x\ t_y\ t_z)^t$ and

$$R = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}$$

R and T describe the orientation and location of the object frame relative to the camera frame respectively. Together, they are referred to as the *camera transformation matrix.* Substituting the parameterized T and R into equation 2 yields

$$\begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix} \tag{3}$$

Combining the projection Eq. 1 with the rigid transformation of Eq. 2 and eliminating λ yields the *collinearity* equations, which describe the ideal relationship between a point on the image plane and the corresponding point in the object frame

$$\begin{aligned} u &= f\frac{r_{11}x + r_{12}y + r_{13}z + t_x}{r_{31}x + r_{32}y + r_{33}z + t_z} \\ v &= f\frac{r_{21}x + r_{22}y + r_{23}z + t_y}{r_{31}x + r_{32}y + r_{33}z + t_z} \end{aligned} \tag{4}$$

For a rigid body transformation, the rotation matrix R must be orthonormal, that is, $R^t = R^{-1}$. The constraint $R^t = R^{-1}$ amounts to the six orthonormality constraint equations on the elements of R

$$\begin{aligned} r_{11}^2 + r_{12}^2 + r_{13}^2 &= 1 \qquad & r_{11}r_{21} + r_{12}r_{22} + r_{13}r_{23} &= 0 \\ r_{21}^2 + r_{22}^2 + r_{23}^2 &= 1 \qquad & r_{11}r_{31} + r_{12}r_{32} + r_{13}r_{33} &= 0 \\ r_{31}^2 + r_{32}^2 + r_{33}^2 &= 1 \qquad & r_{21}r_{31} + r_{22}r_{32} + r_{23}r_{33} &= 0 \end{aligned} \tag{5}$$

where the three constraints on the left column are referred to as the normality constraints and the three on the right column as the orthogonality constraints. The normality constraints ensure that the row vectors of R are unit vectors, while the orthogonality constraints guarantee orthogonality among row vectors.

3. Camera transformation matrix from point correspondences

Given the 3D object coordinates of a number of points and their corresponding 2D image coordinates, the coefficients of R and T can be solved for by a least-squares solution of an over-determined system of linear equations. Specifically, the least-squares method based on point correspondences can be formulated as follows.

Let $X_n = (x_n, y_n, z_n)$, $n = 1, \ldots, K$, be the 3D coordinates of K points relative to the object frame and $U_n = (u_n, v_n)$ be the observed image coordinates of these points. We can then relate X_n and U_n via the collinearity equations in Eq. 4. Rewriting Eq. 4 yields

$$\begin{aligned} fr_{11}x_n + fr_{12}y_n + fr_{13}z_n - u_n r_{31}x_n - u_n r_{32}y_n - u_n r_{33}z_n + ft_x - u_n t_z &= 0 \\ fr_{21}x_n + fr_{22}y_n + fr_{23}z_n - v_n r_{31}x_n - v_n r_{32}y_n - v_n r_{33}z_n + ft_y - v_n t_z &= 0 \end{aligned} \tag{6}$$

We can then set up a matrix M and a vector V as follows

$$M^{2K\times 12} = \begin{pmatrix} fx_1 & fy_1 & fz_1 & 0 & 0 & 0 & -u_1x_1 & -u_1y_1 & -u_1z_1 & f & 0 & -u_1 \\ 0 & 0 & 0 & fx_1 & fy_1 & fz_1 & -v_1x_1 & -v_1y_1 & -v_1z_1 & 0 & f & -v_1 \\ \vdots & & & & & & & & & & & \\ fx_K & fy_K & fz_K & 0 & 0 & 0 & -u_Kx_K & -u_Ky_K & -u_Kz_K & f & 0 & -u_K \\ 0 & 0 & 0 & fx_K & fy_K & fz_K & -v_Kx_K & -v_Ky_K & -v_Kz_K & 0 & f & -v_K \end{pmatrix} \tag{7}$$

$$V^{12\times 1} = \begin{pmatrix} r_{11} & r_{12} & r_{13} & r_{21} & r_{22} & r_{23} & r_{31} & r_{32} & r_{33} & t_x & t_y & t_z \end{pmatrix}^t \tag{8}$$

where M is hereafter referred to as the collinearity matrix and V is the unknown vector of transformation parameters containing all sought rotational and translational coefficients.

To determine V, we can set up a least-squares problem that minimizes

$$\xi^2 = ||MV||^2 \tag{9}$$

where ξ^2 is the sum of squared residual errors of all points. Given an overdetermined system, a least-squares solution to the above equation requires minimization of $||MV||^2$. Its solution contains an arbitrary scale factor due to the lack of constraints on R. To uniquely determine V, different methods have been proposed to solve for the scale factor. In the least-squares solution provided by Sutherland (1974), the depth of the object is assumed to be unity; $t_z = 1$. Not only is this assumption unrealistic for most applications, but also the solution is constructed without regard to the orthonormal constraints that R must satisfy. Faugeras (1993) posed the problem as a constrained least-squares problem using a minimum of 6 points. The third normality constraint (the last one on the left column) in Eq. 5 is imposed by Faugeras during the minimization to solve for the scale factor and to constrain the rotation matrix. The linearity in solution is preserved due to the use of a single normality constraint.

4. Camera transformation matrix from line correspondences

Given correspondences between a set of 3D lines and their observed 2D images, we can set up a system of linear equations that involve R, T, and the coefficients for 3D and 2D lines as follows. Let a 3D line L in the object frame be parametrically represented as

$$L: \quad X = \lambda N + P$$

where $X = (x\ y\ z)^t$ is a generic point on the line, λ is a scalar representing the signed distance from point P to point X, $N = (A\ B\ C)^t$ is the known direction cosine vector

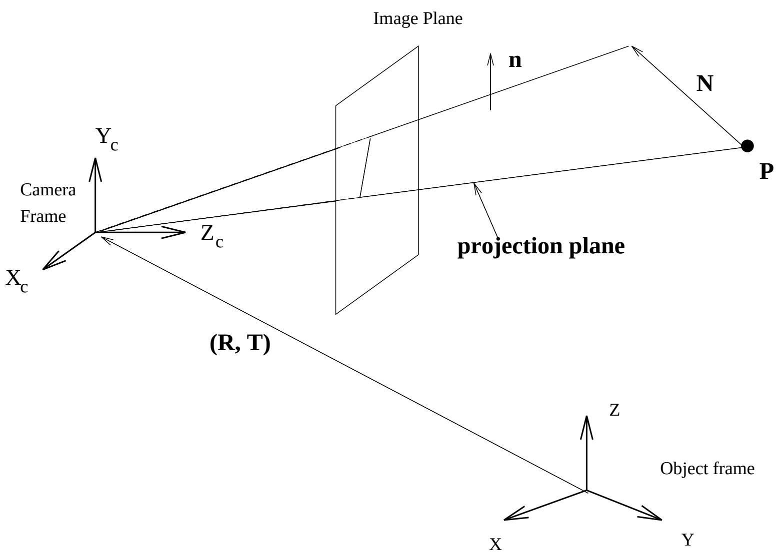

Fig. 2. Projection plane formed by a 2D image line l and the corresponding 3D line L.

and $P = (P_x \; P_y \; P_z)^t$ is a known point on the line relative to the object frame. Let the corresponding 2D line l on the image plane be represented by

$$l: \quad au + bv + c = 0$$

Ideally, the 3D line must lie on the projection plane formed by the center of perspectivity and the 2D image line as shown in Fig. 2.

Relative to the camera frame, the equation of the projection plane can be derived from the 2D line equation as

$$afx_c + bfy_c + cz_c = 0$$

where f is the focal length. Since the 3D line lies on the projection plane, the plane normal must be perpendicular to the line. Denote the plane normal by $n = \frac{(af, bf, c)^t}{\sqrt{a^2f^2+b^2f^2+c^2}}$; then given an ideal projection, we have

$$n^t R N = 0 \tag{10}$$

Similarly, since point P is also located on the projection plane, this leads to

$$n^t(RP + T) = 0 \tag{11}$$

Eqs. 10 and 11 are hereafter referred to as *coplanarity* equations. Equivalently, they can be rewritten as

$$A\ a\ r_{11} + B\ a\ r_{12} + C\ a\ r_{13} + A\ b\ r_{21} + B\ b\ r_{22} + C\ b\ r_{23} + A\ c\ r_{31} + B\ c\ r_{32} + C\ c\ r_{33} = 0$$
$$P_x\ a\ r_{11} + P_y\ a\ r_{12} + P_z\ a\ r_{13} + P_x\ b\ r_{21} + P_y\ b\ r_{22} + P_z\ b\ r_{23} + P_x\ c\ r_{31} + P_y\ c\ r_{32} + P_z\ c\ r_{33} + a\ t_x + b\ t_y + c\ t_z = 0$$

Given a set of J line correspondences, we can set up a system of linear equations similar to those for points that involve matrix H and vector V, where V is as defined before and H is defined as follows

$$H^{2J\times 12} = \begin{pmatrix} A_1a_1 & B_1a_1 & C_1a_1 & A_1b_1 & B_1b_1 & C_1b_1 & A_1c_1 & B_1c_1 & C_1c_1 & 0 & 0 & 0 \\ P_{x_1}a_1 & P_{y_1}a_1 & P_{z_1}a_1 & P_{x_1}b_1 & P_{y_1}b_1 & P_{z_1}b_1 & P_{x_1}b_1 & P_{y_1}c_1 & P_{z_1}c_1 & a_1 & b_1 & c_1 \\ \vdots & & & & & & & & & & & \\ A_Ja_J & B_Ja_J & C_Ja_J & A_Jb_J & B_Jb_J & C_Jb_J & A_Jc_J & B_Jc_J & C_Jc_J & 0 & 0 & 0 \\ P_{x_J}a_J & P_{y_J}a_J & P_{z_J}a_J & P_{x_J}b_J & P_{y_J}b_J & P_{z_J}b_J & P_{x_J}c_J & P_{y_J}c_J & P_{z_J}c_J & a_J & b_J & c_J \end{pmatrix} \quad (12)$$

and is called the coplanarity matrix. Again we can solve for V by minimizing the sum of squared residual errors $||HV||^2$. V can be solved for up to a scale factor. The scale factor can be determined by imposing one of the normality constraints as discussed for the case using points.

5. Camera transformation matrix from ellipse-circle correspondences

5.1. *Camera transformation matrix from circles*

Given the image (an ellipse) of a 3D circle and its size, the pose (position and orientation) of the 3D circle relative to the camera frame can be solved for analytically. Solutions to this problem may be found in Haralick and Shapiro (1993), Forsyth *et al.* (1991), and Dhome *et al.* (1989). If we are also given the pose of the circle in the object frame, then we can use the two poses to solve for R and T. Specifically, let $N_c = (N_{c_x}\ N_{c_y}\ N_{c_z})^t$ and $O_c = (O_{c_x}\ O_{c_y}\ O_{c_z})^t$ be the 3D circle normal and center in camera coordinate frame respectively. Also, let $N_o = (N_{o_x}\ N_{o_y}\ N_{o_z})^t$ and $O_o = (O_{o_x}\ O_{o_y}\ O_{o_z})^t$ be the normal and center of the same circle, but in the object coordinate system. O_c and N_c are computed from the observed image ellipse using a technique described in Forsyth *et al.* (1991) while N_o and O_o are assumed to be known. The problem is to determine R and T from the correspondence between N_c and N_o, and between O_c and O_o. The two normals and the two centers are related by the transformation R and T as shown below

$$\begin{aligned} N_c &= RN_o \\ &= \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \begin{pmatrix} N_{o_x} \\ N_{o_y} \\ N_{o_z} \end{pmatrix} \end{aligned} \quad (13)$$

and

$$\begin{aligned} O_c &= R\,O_o + T \\ &= \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \begin{pmatrix} O_{o_x} \\ O_{o_y} \\ O_{o_z} \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix} \end{aligned} \quad (14)$$

Equivalently, we can rewrite Eqs. 13 and 14 as follows

$$\begin{aligned} N_{o_x}r_{11} + N_{o_y}r_{12} + N_{o_z}r_{13} &= N_{c_x} \\ N_{o_x}r_{21} + N_{o_y}r_{22} + N_{o_z}r_{23} &= N_{c_y} \\ N_{o_x}r_{31} + N_{o_y}r_{32} + N_{o_z}r_{33} &= N_{c_z} \end{aligned}$$

and

$$
\begin{aligned}
O_{o_x} r_{11} + O_{o_y} r_{12} + O_{o_z} r_{13} + t_x &= O_{c_x} \\
O_{o_x} r_{21} + O_{o_y} r_{22} + O_{o_z} r_{23} + t_y &= O_{c_y} \\
O_{o_x} r_{31} + O_{o_y} r_{32} + O_{o_z} r_{33} + t_z &= O_{c_z}
\end{aligned}
$$

Each pair of 2D ellipse and 3D circle therefore offers 6 equations. The 3 equations from orientation (Eq. 13) are not independent due to unity constraint on the normals. Given I observed ellipses and their corresponding object space circles, we can set up a system of linear equations to solve for R and T by minimizing the sum of residual errors $||QV - k||^2$, where Q and k are defined as follows

$$
Q^{6I\times 12} = \begin{pmatrix}
N_{1_{o_x}} & N_{1_{o_y}} & N_{1_{o_z}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & N_{1_{o_x}} & N_{1_{o_y}} & N_{1_{o_z}} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & N_{1_{o_x}} & N_{1_{o_y}} & N_{1_{o_z}} & 0 & 0 & 0 \\
O_{1_{o_x}} & O_{1_{o_y}} & O_{1_{o_z}} & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & O_{1_{o_x}} & O_{1_{o_y}} & O_{1_{o_z}} & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & O_{1_{o_x}} & O_{1_{o_y}} & O_{1_{o_z}} & 0 & 0 & 1 \\
\vdots & & & & & & & & & & & \\
N_{I_{o_x}} & N_{I_{o_y}} & N_{I_{o_z}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & N_{I_{o_x}} & N_{I_{o_y}} & N_{I_{o_z}} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & N_{I_{o_x}} & N_{I_{o_y}} & N_{I_{o_z}} & 0 & 0 & 0 \\
O_{I_{o_x}} & O_{I_{o_y}} & O_{I_{o_z}} & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & O_{I_{o_x}} & O_{I_{o_y}} & O_{I_{o_z}} & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & O_{I_{o_x}} & O_{I_{o_y}} & O_{I_{o_z}} & 0 & 0 & 1
\end{pmatrix} \tag{15}
$$

and

$$
k^{6I\times 1} = (N_{1_{c_x}}\ N_{1_{c_y}}\ N_{1_{c_z}}\ O_{1_{c_x}}\ O_{1_{c_y}}\ O_{1_{c_z}}\ \ldots\ N_{I_{c_x}}\ N_{I_{c_y}}\ N_{I_{c_z}}\ O_{I_{c_x}}\ O_{I_{c_y}}\ O_{I_{c_z}})^t \tag{16}
$$

Since each circle provides 6 equations, a minimum of 2 circles are needed (if only circles are used) to uniquely solve for the 12 parameters in the transformation matrix. To retain a linear solution, not even one normality constraint can be imposed using Lagrange multipliers due to k being non-zero vector.

6. The integrated technique

In the previous sections we have outlined the least-squares frameworks for computing the transformation matrix from different features individually. It is desirable to be able to compute camera exterior parameters using more than one type of feature *simultaneously*. In other words, given observed geometric entities at different levels, we want to develop a mechanism that systematically and consistently fuses this information. The reason is quite obvious: using all available geometric information will provide a more accurate and robust solution, since it increases the redundancy of the least-squares estimation. It also reduces the dependency on points. We can be more selective when choosing points, without worrying about the minimum number of points needed for accurate results. Furthermore, we can be less worry about whether the selected points are coplanar or not and their distribution. This section is devoted to formulating a direct solution for computing camera transformation matrix from the simultaneous use of 2D-3D correspondences of points, lines, and ellipses/circles.

6.1. *Fusing all observed information*

The problem of integrating information from points, lines, and circles is actually straightforward, given the frameworks we have outlined individually for points, lines, and circles. The problem can be stated as follows

Given the 2D-3D correspondences of K points, J lines, and I ellipse/circle pairs, we want to set up a system of linear equations that involves all geometric entities. The problem can be formulated as a least-squares estimation in the form of minimizing $||WV - b||$, where V is the unknown vector of transformation parameters as defined before, and b is a known vector defined below. W is an augmented coefficient matrix, whose rows consist of linear equations derived from points, lines, and circles. Specifically, given the M, H, and Q matrices defined in Eqs. 7,12, and 15, the W matrix is

$$W^{(2K+2J+6I)\times 12} = \begin{pmatrix} M \\ H \\ Q \end{pmatrix} \tag{17}$$

where the first $2K$ rows of W represent contributions from points, the second subsequent $2J$ rows represent contributions from lines, and the last $6I$ rows represent contributions from circles. The vector b is defined as

$$b^{(2K+2J+6I)\times 1} = (0 \;\; 0 \;\; 0 \;\; \ldots \;\; k)^t \tag{18}$$

where k is defined in Eq. 16. Given W and b, the least-squares solution to V is

$$V = (W^t W)^{-1} W^t b \tag{19}$$

It can be seen that to have an overdetermined system of linear equations, we need $2K + 2J + 6I \geq 12$ observed geometric entities. This may occur with any combination of points, lines, and circles. For example, one point, one line, and one circle or two points and one circle are sufficient to solve the transformation matrix from Eq. 19. Any additional points or lines or circles will improve the robustness and the precision of the estimated parameters. The integrated approach so far, however, does not impose any of the six orthonormality constraints. Also, we assume equal weightings for observations from the multiple features. We realize stochastic model is very important while using multiple features. As part of future work, we plan to use error propagation to compute a covariance matrix for each type of feature. The covariance matrices can then be used to compute the observation weightings.

6.2. *Approximately imposing orthonormal constraints*

The least-squares solution to V described in the last section cannot guarantee the orthonormality of the resultant rotation matrix. One major reason why previous linear methods are very susceptible to noise is because the orthonormality constraints are not enforced or enforced weakly. To ensure this, the six orthonormal constraints must be imposed on V within the least-squares framework. The conventional way of imposing these constraints is through the use of Lagrange multipliers. However, simultaneously imposing any two normality constraints or one orthogonality constraint using Lagrange multipliers requires a non-linear solution for the problem. Therefore, most linear methods choose to use a single normality constraint. For example, Faugeras (1993) imposed the constraint that the norm of the last row vector of R be unity. This constraint, however, cannot ensure complete satisfaction of all orthonormal constraints. To impose more than one orthonormality constraint but still retain a linear solution, Liu *et al.* (1990) suggested the constraint that the sum of the squares of the three row vectors be 3. This constraint, however, cannot guarantee normality of each individual row vector. Haralick *et al.* (1989) and Horn (1987) proposed direct solutions, where all orthonormality constraints are imposed, but for the 3D to 3D absolute orientation problem. They are only applicable for point correspondences and are not applicable to line and circle-ellipse correspondences. Most important, their techniques cannot be applied to the general linear framework we propose.

Given the linear framework, even imposing one constraint using the Lagrange multiplier can render the solution non-linear. It is well-known that the solution to minimizing a quadratic function with quadratic constraints (referred to as *trust region* problem in statistics) can only be achieved by non-linear methods.

We now introduce a simple, yet effective, method for approximately imposing the orthonormal constraints in a way that offers a linear solution. We want to emphasize that the technique we are about to introduce cannot guarantee a perfect (satisfying $R^t = R^{-1}$) rotation matrix. However, our experimental study proves that it yields a matrix that is closer to a rotation matrix than those obtained using the competing methods. The advantages of our technique are: 1) all six orthonormal constraints are imposed simultaneously; 2) constraints are imposed on each entry of the rotation matrix, and (3) asymptotically, the resulting matrix should converge to a rotation matrix. However, this method requires knowledge of the 3D orientation of a vector in both camera and object frames.

We now address the problem of how to impose the orthonormality constraints in the general framework of finding pose from multiple geometric features described in sections 3, 4, and 5. Given the pose of circles relative to the camera frame and the object frame, let $N_c = (N_{c_x}\ N_{c_y}\ N_{c_z})^t$ and $N_o = (N_{o_x}\ N_{o_y}\ N_{o_z})^t$ be the 3D circle normals in camera and object frames respectively. Eq. 13 depicts the relation between two normals that involves R. The relation can also be expressed in an alternative way that involves R^t (note $R^t = R^{-1}$) as follows

$$\begin{aligned} N_o &= R^t N_c \\ &= \begin{pmatrix} r_{11} & r_{21} & r_{31} \\ r_{12} & r_{22} & r_{32} \\ r_{13} & r_{23} & r_{33} \end{pmatrix} \begin{pmatrix} N_{c_x} \\ N_{c_y} \\ N_{c_z} \end{pmatrix} \end{aligned} \tag{20}$$

Equivalently, we can rewrite Eq. 20 as follows

$$\begin{aligned} N_{c_x} r_{11} + N_{c_y} r_{21} + N_{c_z} r_{31} &= N_{o_x} \\ N_{c_x} r_{12} + N_{c_y} r_{22} + N_{c_z} r_{32} &= N_{o_y} \\ N_{c_x} r_{13} + N_{c_y} r_{23} + N_{c_z} r_{33} &= N_{o_z} \end{aligned}$$

Given the same set of I observed ellipses and their corresponding object space circles, we can set up another system of linear equations that uses the same set of circles as in Q. Let Q' be the coefficient matrix that contains the coefficients of the set of linear equations, then Q' is

$$Q'^{3I\times 12} = \begin{pmatrix} N_{1_{cx}} & 0 & 0 & N_{1_{cy}} & 0 & 0 & N_{1_{cz}} & 0 & 0 & 0 & 0 & 0 \\ 0 & N_{1_{cx}} & 0 & 0 & N_{1_{cy}} & 0 & 0 & N_{1_{cz}} & 0 & 0 & 0 & 0 \\ 0 & 0 & N_{1_{cx}} & 0 & 0 & N_{1_{cy}} & 0 & 0 & N_{1_{cz}} & 0 & 0 & 0 \\ \vdots & & & & & & & & & & & \\ N_{I_{cx}} & 0 & 0 & N_{I_{cy}} & 0 & 0 & N_{I_{cz}} & 0 & 0 & 0 & 0 & 0 \\ 0 & N_{I_{cx}} & 0 & 0 & N_{I_{cy}} & 0 & 0 & N_{I_{cz}} & 0 & 0 & 0 & 0 \\ 0 & 0 & N_{I_{cx}} & 0 & 0 & N_{I_{cy}} & 0 & 0 & N_{I_{cz}} & 0 & 0 & 0 \end{pmatrix} \tag{21}$$

Correspondingly, we have k' defined as

$$k'^{3I\times 1} = (N_{1_{ox}}\ N_{1_{oy}}\ N_{1_{oz}}\ 0\ 0\ 0 \ldots\ N_{I_{ox}}\ N_{I_{oy}}\ N_{I_{oz}}\ 0\ 0\ 0)^t \tag{22}$$

To implement the constraint in the least-squares framework, we can augment matrix W in Eq. 19 with Q', yielding W', and augment vector b in Eq. 18 with k', yielding b', where W' and b' are defined as follows

$$W'^{(2K+2J+9I)\times 12} = \begin{pmatrix} W \\ Q' \end{pmatrix} \qquad b'^{(2K+2J+9I)\times 1} = \begin{pmatrix} b \\ k' \end{pmatrix}$$

Putting it all together, the solution to V can be found by minimizing $||W'V - b'||^2$, given by

$$V = (W'^t W')^{-1} W'^t b' \tag{23}$$

For each circle, we add three additional equations to impose the orthonormality constraints. Therefore, the first $2K$ rows of W' represent contributions from points, the second subsequent $2J$ rows represent contributions from lines, and the last $9I$ rows represent contributions from circles.

We must make it clear that the proposed method for imposing orthonormality will not be applicable if the required ellipse/circle features are not available. It is only applicable to circles but not applicable to points or lines. In this case, we can only partially apply the orthonormality constraints, i.e., applying one of the normality constraints or performing a post-orthogonalization process like the one used by Faugeras (1993).

To have an overdetermined system of linear equations, we need $2K + 2J + 9I \geq 12$ observed geometric entities. This may occur with any combination of points, lines, and circles. For example, one point, one line, and one circle or two points and one circle or even two circles are sufficient to solve for the transformation matrix. The resultant transformation parameters R and T are more accurate and robust due to fusing information from different sources. The resultant rotation matrix R is also very close to being orthonormal since the orthonormality constraints have been explicitly added to the system of linear equations used in the least-squares estimation. The obtained R and T can be used directly for certain applications or fed to an iterative procedure (such as the bundle adjustment method) to refine the solution. Since the obtained transformation parameters are accurate, the subsequent iterative procedure not only can converge very quickly, usually after a couple of iterations as evidenced by our experiments but, more importantly, converge correctly . In the section to follow, we study the performance of the new linear exterior orientation estimation method against the methods that use only one type of geometric entity at a time, using synthetic and real images.

7. Experiments

In this section, we present and discuss the results of a series of experiments aimed at characterizing the performance of the integrated linear exterior orientation estimation technique. Using both synthetic data and real images of industrial parts, the experiments conducted aim at studying the effectiveness of the proposed technique for approximately imposing the orthonormal constraints and quantitatively evaluating the overall performance of the integrated linear technique against existing linear techniques.

7.1. *Performance study with synthetic data*

This section consists of two parts. First, we present results from a large number of controlled experiments aimed at analyzing the effectiveness of our technique for imposing orthonormal constraints. This was accomplished by comparing the errors of the estimated rotation and translation vectors obtained with and without orthonormal constraints imposed under different conditions. Second, we discuss the results from a comparative performance study of the integrated linear technique against an existing linear technique under different noisy conditions.

In the experiments with simulation data, the 3D data (3D point coordinates, 3D surface normals, 3D line direction cosines) are generated randomly within specified ranges. For example, 3D coordinates are randomly generated within the cube [(-5, -5, -5) to (5,5,5)]. 2D data are generated by projecting the 3D data onto the image plane, followed by perturbing the projected image data with independently and identically distributed (iid) Gaussian noise of mean vector of zero and covariance matrix of σ^2I, where σ^2 represents noise variance and I is the identity matrix. From the generated 2D-3D data, we estimate the rotation and the translation vector using the linear algorithm, from which we can compute the estimation errors. The estimation error is defined as the Euclidean distance between the estimated rotation (translation) vector and the ideal rotation (translation) vector. We choose the rotation matrix rather than other specific representations like Euler angles and quaternion parameters for error analysis. This is because all other representations depend on the estimated rotation vector. For each experiment, 100 trials (with different noise instances) are performed and the average distance errors are computed. The noise level is quantified using the signal to noise ratio (SNR). SNR is defined as $20log\frac{d}{\sigma}$, where σ is the standard deviation of the Gaussian noise and d is the range of the quantity being perturbed.

Figs. 3 and 4 plot the mean rotation and translation errors as a function of the signal to noise ratio, with and without orthonormal constraints imposed. It is clear from the two figures that imposing the orthonormal constraints improves the estimation errors for both the rotation and translation vectors. The improvement is especially significant when the SNR is low.

To further study the effectiveness of the technique for imposing constraints, we studied

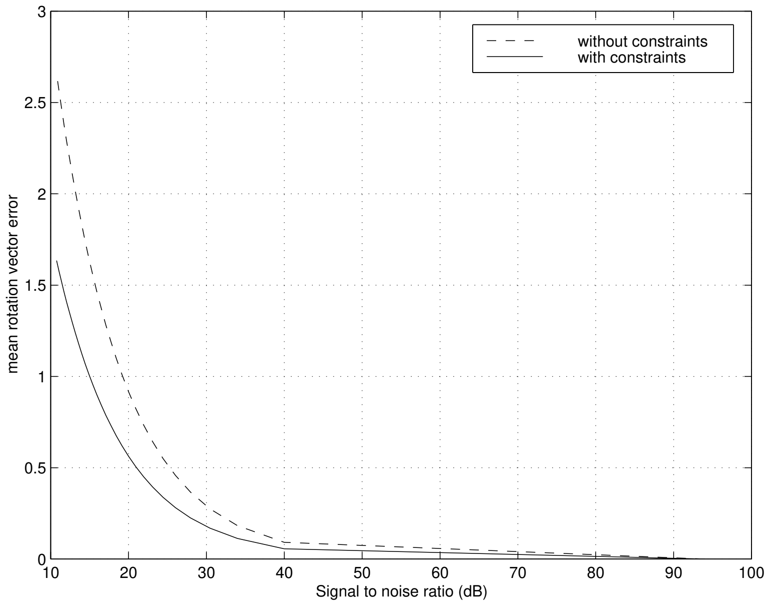

Fig. 3. Mean rotation vector error versus SNR. The plot was generated using the integrated linear technique with a combination of one point, one line and one circle. Each point in the figure represents an average of 100 trials.

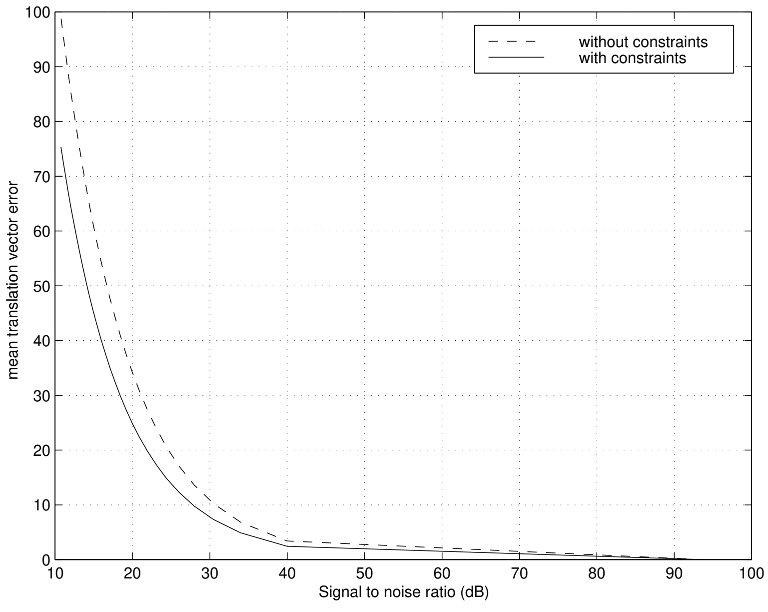

Fig. 4. Mean translation vector error (mm) versus SNR. The plot was generated using the integrated linear technique with a combination of one point, one line and one circle. Each point in the figure represents an average of 100 trials.

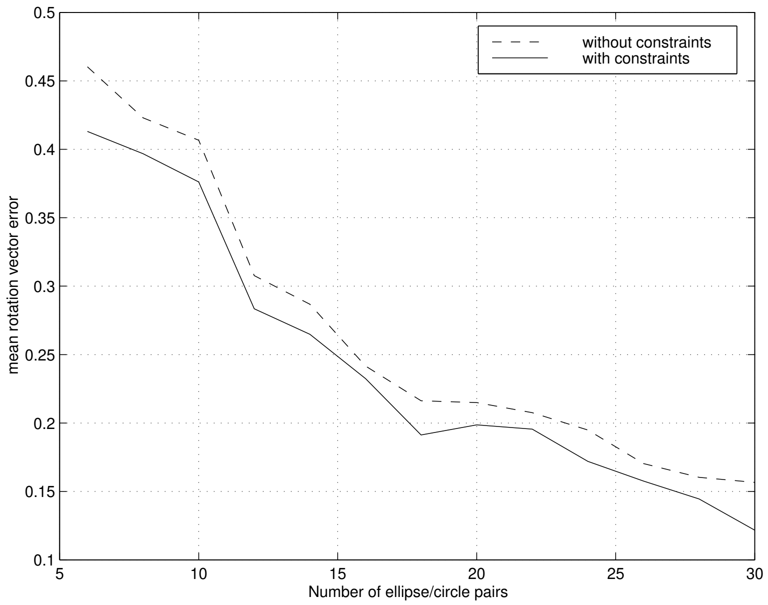

Fig. 5. Mean rotation vector error versus the number of ellipse/circle pairs (SNR=35)

its performance under different numbers of pairs of ellipse/circle correspondences. This experiment is intended to study the efficacy of imposing orthonormal constraints versus the amount of geometric data used for the least-squares estimation. The results are plotted in Figs. 5 and 6, which give the average rotation and translation errors as a function of the number of ellipse/circle pairs used, with and without constraints imposed. The two figures again show that imposing orthonormal constraints leads to an improvement in estimation errors. This improvement, however, begins to decrease when the features used exceed a certain number. The technique is most effective when fewer ellipse/circle pairs are used.

To compare the integrated linear technique with an existing linear technique, we studied its performance against that of Faugeras (1993). The results are given in Figs. 7 and 8, which plot the mean rotation and translation vector errors as a function of the SNR respectively.

The two figures clearly show the superiority of the new integrated technique over Faugeras' linear technique, especially for the translation errors. To further compare the sensitivity of the two techniques to viewing parameters, we changed the position parameters of the camera by doubling the camera distance z to the object points. Figs. 9 and 10 plot the mean rotation and translation vector errors as a function of SNR respectively under the new camera position. While increasing z causes an increase in the estimation errors for both techniques, its impact on Faugeras' technique is more serious. This leads to a much more noticeable performance difference between the two linear techniques. The fact that the integrated technique using only five geometric entities (3 points, 1 line, and 1 circle with a total of 17 equations) still outperforms Faugeras' technique, which uses 10

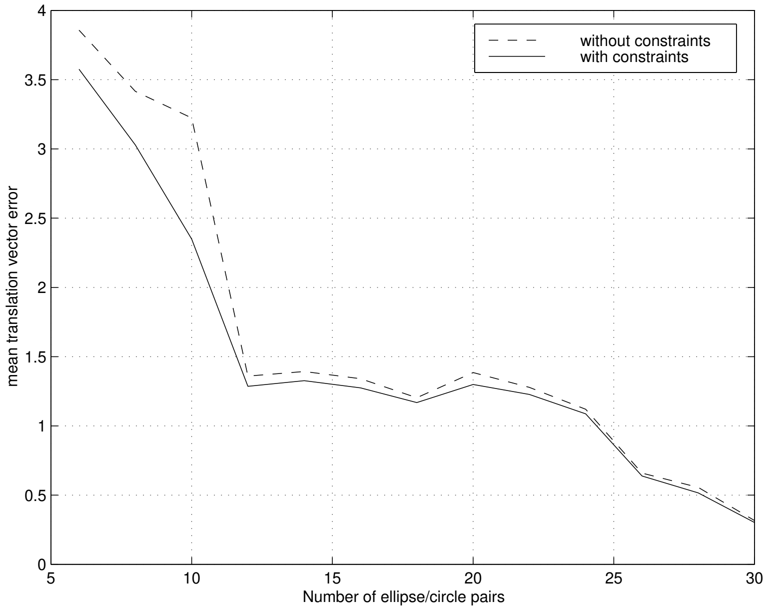

Fig. 6. Mean translation vector error (mm) versus the number of ellipse/circle pairs (SNR=35)

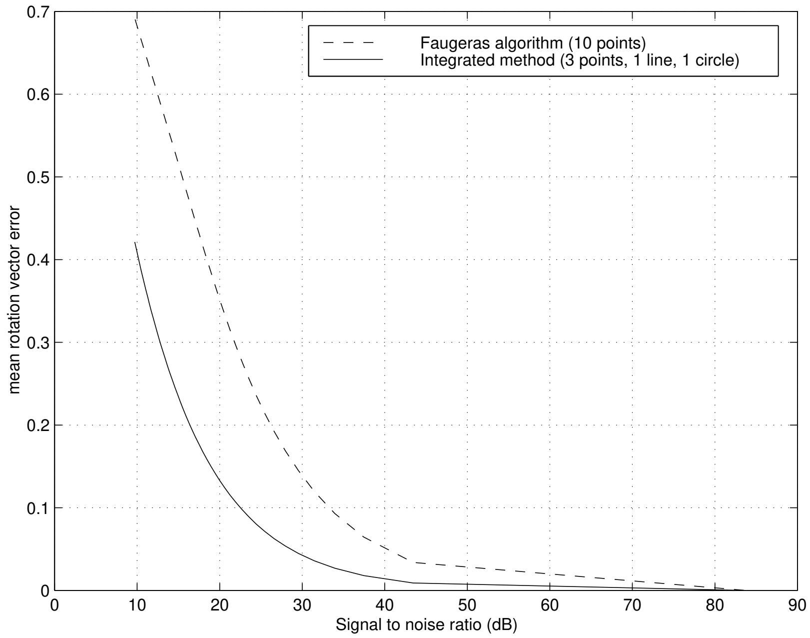

Fig. 7. Mean rotation vector error versus SNR. The curve for Faugeras' algorithm was obtained using 10 points, while the curve for the integrated technique was generated using a combination of three points, one line, and one circle.

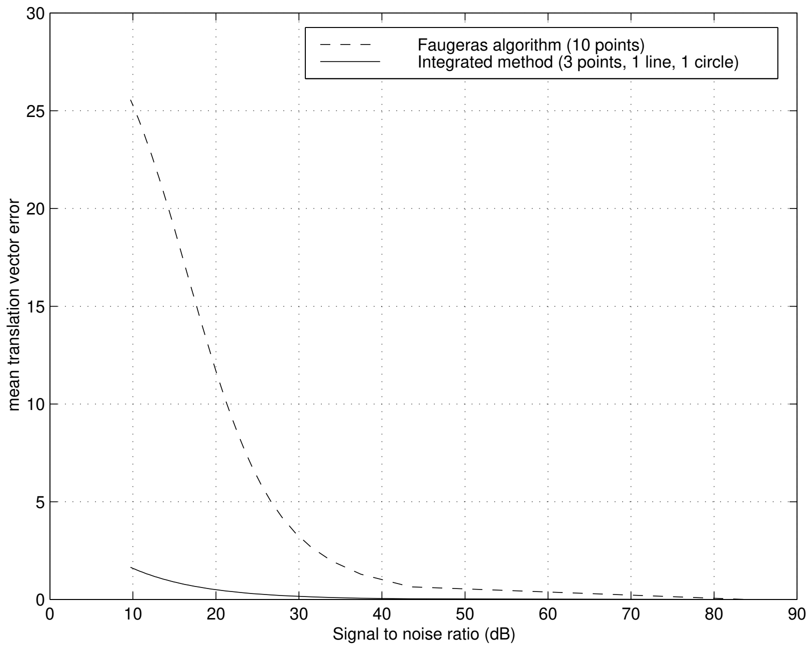

Fig. 8. Mean translation vector error (mm) versus SNR. The curve for Faugeras' algorithm was obtained using 10 points, while the curve for the integrated technique was generated using a combination of three points, one line, and one circle.

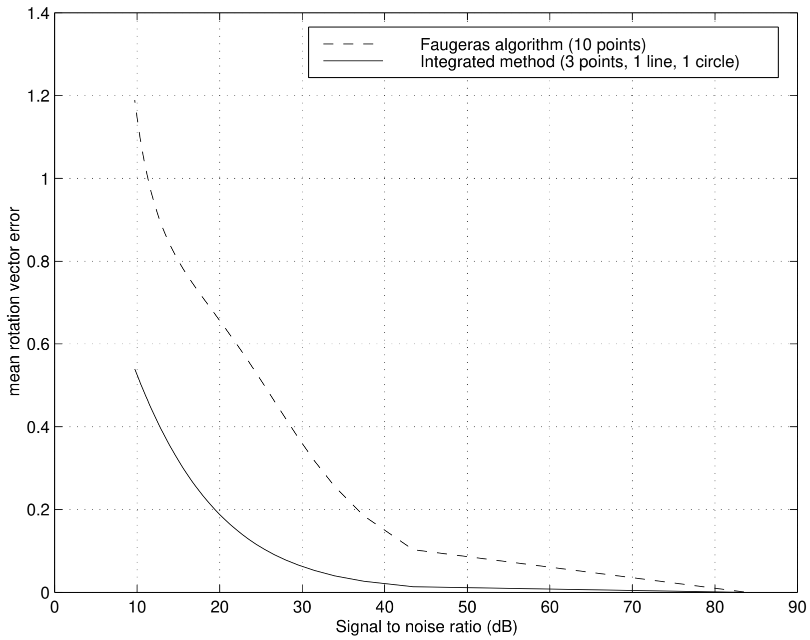

Fig. 9. Mean rotation vector error versus SNR with an increased camera position parameter z (doubling the camera and object distance z). The curve for Faugeras' algorithm was obtained using 10 points, while the curve for the integrated technique was generated using a combination of three points, one line, and one circle.

points (a total of 20 equations), shows that the higher-level geometric features such as lines and circles can provide more robust solutions than those provided solely by points. This demonstrates the power of combining features on different levels of abstraction. Our study also shows that Faugeras' linear technique is very sensitive to noise when the number of points used is close to the required minimum. For example, when only six points are used, a small perturbation of the input data can cause significant errors on the estimated parameters, especially the translation vector. Figs. 9 and 10 reveal that the technique using only points is numerically unstable to viewing parameters with respect to z.

7.2. *Performance characterization with real images*

This section presents results obtained using real images of industrial parts. The images contain linear features (points and lines) and non-linear features (circles). This phase of the experiments consist of two parts. First, we visually assess the performance of the linear least-squares framework using different combinations of geometric entities, such as one circle and six points; one circle and two points; and one circle, one point and two lines. The performance of the proposed technique is judged by visual inspection of the alignment between the image of a part and the reprojected outline of the part using the estimated transformation matrix. Second, the technique is compared against existing methods that use only one type of geometric entity as well as against the Gauss-Newton

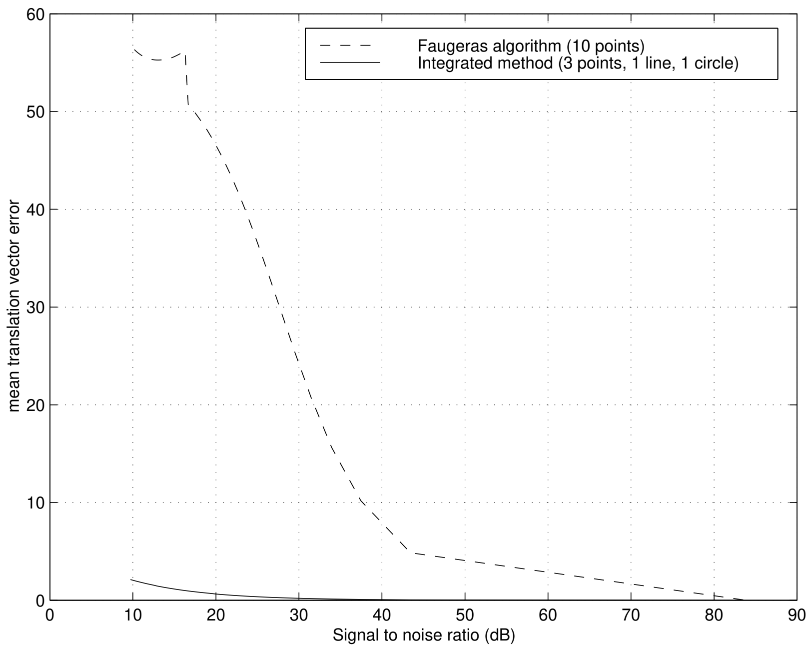

Fig. 10. Mean translation vector error (mm) versus SNR with an increased camera position parameter z (doubling the camera and object distance z). The curve for Faugeras' algorithm was generated using 10 points while the curve for the integrated technique was generated using a combination of three points, one line, and one circle.

iterative method (similar to the bundle adjustment method). Due to the unavailability of the groundtruth data, the closeness between the solutions from the integrated linear method and from the iterative method, as represented by the residual errors, as well as the number of iterations required by the iterative method to converge, are used as measures to indicate the goodness of the solution obtained using the new method.

To test the integrated technique, we performed the following experiments. The transformation ($R\ T$) was analytically computed using our integrated technique with different combinations of points, lines, and circles. Figs. 11, 12, and 13 illustrate the results obtained using the following combinations of features: two points plus one circle; one point, two lines, and one circle; and six points and one circle respectively. Visual inspection of the figures reveals that results obtained from the three configurations are all good enough to serve as an initial guess to an iterative procedure. It is also evident from Figs. 11, 12, and 13 that the result using six points and one circle is superior to the ones obtained using the other two configurations.

The significance of these sample results is as follows. First, they demonstrate the feasibility of the proposed framework applied to real image data. Second, they show that using multiple geometric primitives simultaneously to compute the exterior orientation reduces the dependency on points. One can be more selective when choosing which point correspondence to use in exterior orientation estimation. This can potentially improve the robustness of the estimation procedure since image points are more susceptible to noise than image lines and ellipses. Third, the use of more than the minimum required number of geometric features provides redundancy to the least-squares estimation, therefore improving the accuracy of the solution, as evidenced by the progressively improved results as the number of linear equations increase.

In order to compare the results with those of other existing techniques, we computed the exterior orientation of the same object using the same six points and the same circle, separately. The result for the exterior orientation computation using a linear technique (Ji and Costa, 1997) (similar to that of Faugeras (1993)) with six points is given in Fig. 14. The algorithms of Dhome *et al.* (1989) and Forsyth (Forsyth *et al.*, 1991) for the pose-from-circle computation were augmented in (Costa, 1997) to handle non-rotationally symmetric objects. The results of this augmented algorithm using the single ellipse/circle correspondence is shown in Fig. 15. Notice that due to the localized concentration of detectable feature points and the physical distance between the circle and these points, the projected outlines computed align well only in the areas where the features used are located. Specifically, the result in Fig. 15 shows a good alignment in the upper portion of the object where the circle is located and a poor alignment in the lower part (as indicated by the arrow). On the other hand, the result in Fig. 14 shows a good alignment only at the lower part of the object where the concentration of detectable feature points is located and a poor alignment on the upper part of the object (as indicated by the arrow).

Visual inspection of the results in Figs. 13, 14, and 15 shows the benefits of the new technique over the existing methods. The model reprojection using the transformation matrix obtained using the new technique yields a better alignment than those using only points or only ellipse/circle correspondences. To compare the performance quantitatively, we compare the transformation matrices obtained using the three methods (with different combinations of features) against the one obtained from the iterative procedure (bundle

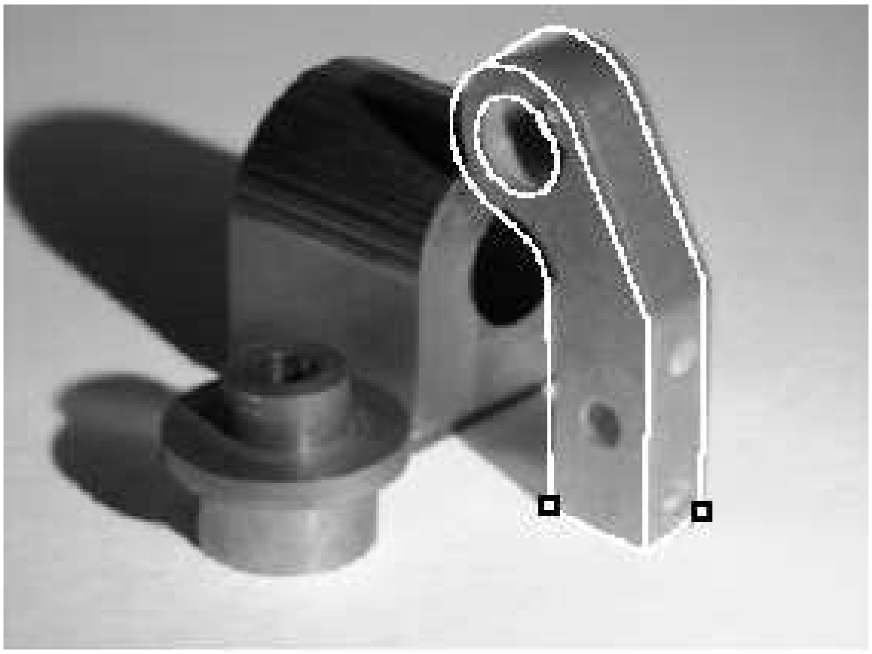

Fig. 11. The alignment between the image of a part and the reprojected outline (white) of using the exterior orientation computed using two points (as indicated by the black squares) and one circle (the upper circle) with the integrated technique.

adjustment) Table 1 shows the numerical results for the transformations obtained from using only points, only the circle, and a combination of points and circle. The results from each method were then used as the initial guess to the iterative Gauss-Newton method. The final transformation obtained after the convergence of the iterative method is shown in the last row of Table 1. These final results are the same regardless of which initial guess was used. But they vary in number of iterations required as shown in table 2.

Table 2 summarizes the number of iterations required for the iterative procedure to converge using as initial guesses the results from the three linear methods mentioned in Table 1. Fig. 16 shows the results from the iterative procedure.

It is evident from Tables 1, 2 and Fig. 16 that the new technique yields a transformation matrix that is closer to the one obtained from the bundle adjustment procedure and therefore requires fewer iterations (one here) for the bundle adjustment method to converge. By contrast, the final results for initial guesses obtained using only points and only one circle require 4 and 6 iterations, respectively, for the bundle adjustment procedure to converge. The result from the quantitative study echoes the conclusion from visual inspection: the new technique offers better estimation accuracy, because it is capable of fusing all information available.

To further validate our technique, we tested it on over fifty real images with similar results. Fig. 17 give results of applying the integrated technique to different industrial parts with different combinations of geometric features. Our experiments also reveal, as was theoretically expected for a system of linear equations of the form $AX = b$, a decay

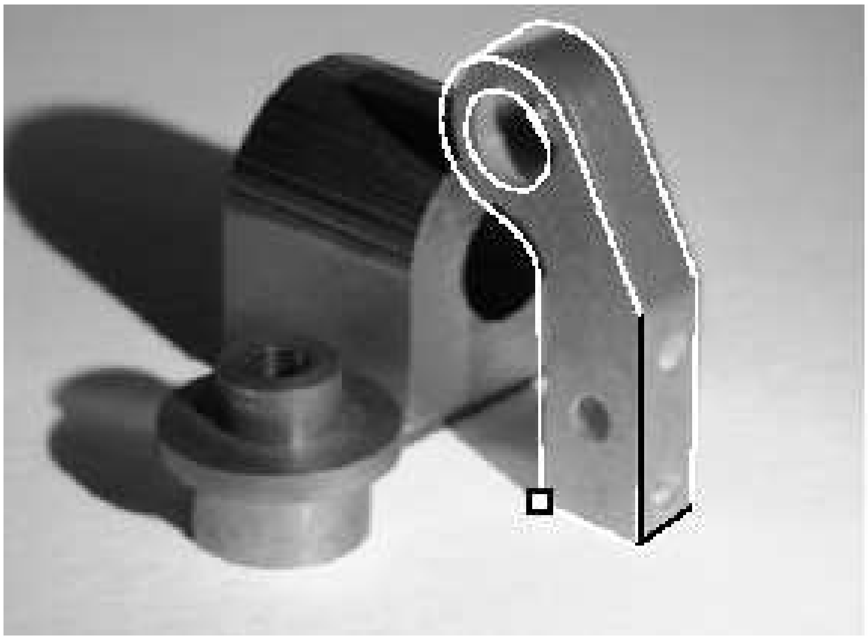

Fig. 12. The alignment between the image of a part and the reprojected outline (white) of using the exterior orientation computed using one point, two lines, and one circle (the upper circle) with the integrated technique. The points and lines used are marked with black squares and lines.

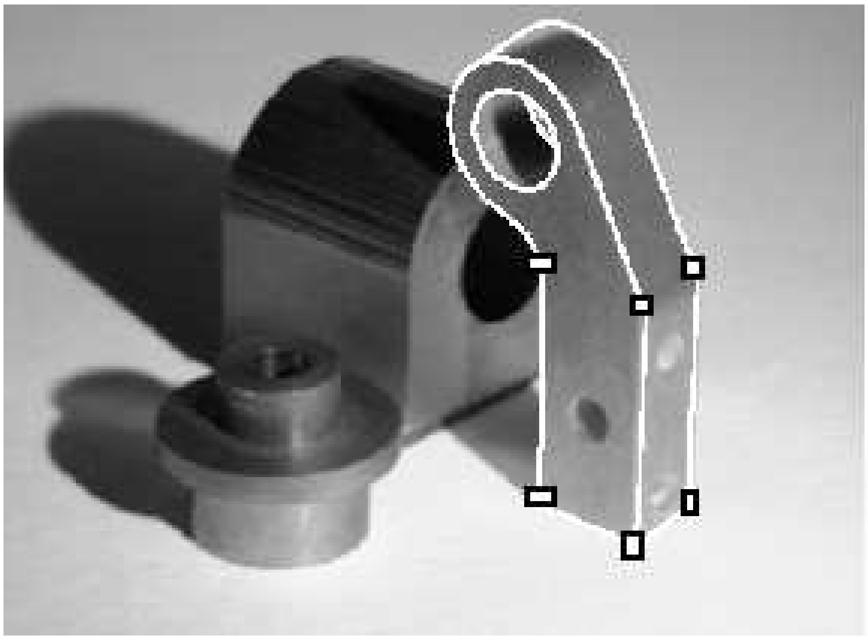

Fig. 13. The alignment between the image of a part and the reprojected outline (white) of using the exterior orientation computed using six points (as indicated by the black squares) and one circle (the upper circle) with the integrated technique.

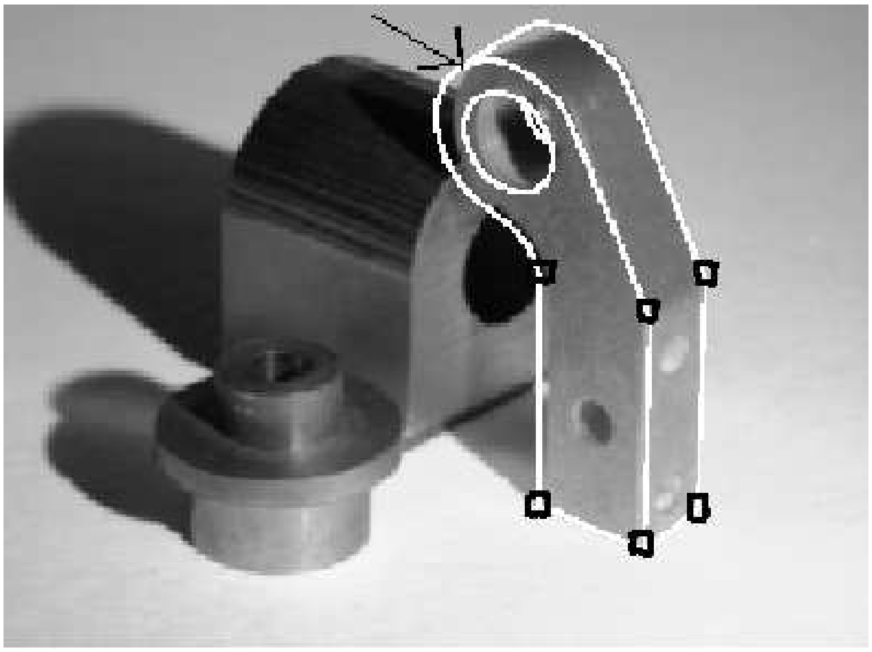

Fig. 14. The alignment between the image of a part and the reprojected outline (white) of using the exterior orientation computed using six points alone. It shows good alignment only at the lower part of the object where the concentration of detectable feature points is located and a poor alignment on the upper part of the object (as indicated by the arrow).

Table 1
Exterior orientations from different methods

Method	R	T (mm)
point only (Fig. 14)	$\begin{bmatrix} 0.410 & -0.129 & -0.902 \\ 0.606 & -0.700 & 0.376 \\ -0.681 & -0.701 & -0.208 \end{bmatrix}$	$[-43.125\ \ -25.511\ 1232.036]$
circle only (Fig. 15)	$\begin{bmatrix} 0.302 & 0.302 & -0.932 \\ 0.692 & -0.628 & 0.355 \\ -0.655 & -0.753 & -0.054 \end{bmatrix}$	$[-35.161\ \ -15.358\ 1195.293]$
points and circle (Fig. 13)	$\begin{bmatrix} 0.398 & -0.142 & -0.902 \\ 0.554 & -0.667 & 0.336 \\ -0.700 & -0.684 & -0.201 \end{bmatrix}$	$[-43.077\ \ -26.400\ 1217.855]$
bundle adjustment (Fig. 16)	$\begin{bmatrix} 0.341 & -0.156 & -0.927 \\ 0.631 & -0.693 & 0.349 \\ -0.697 & -0.704 & -0.137 \end{bmatrix}$	$[-43.23\ \ -28.254\ 1273.07]$

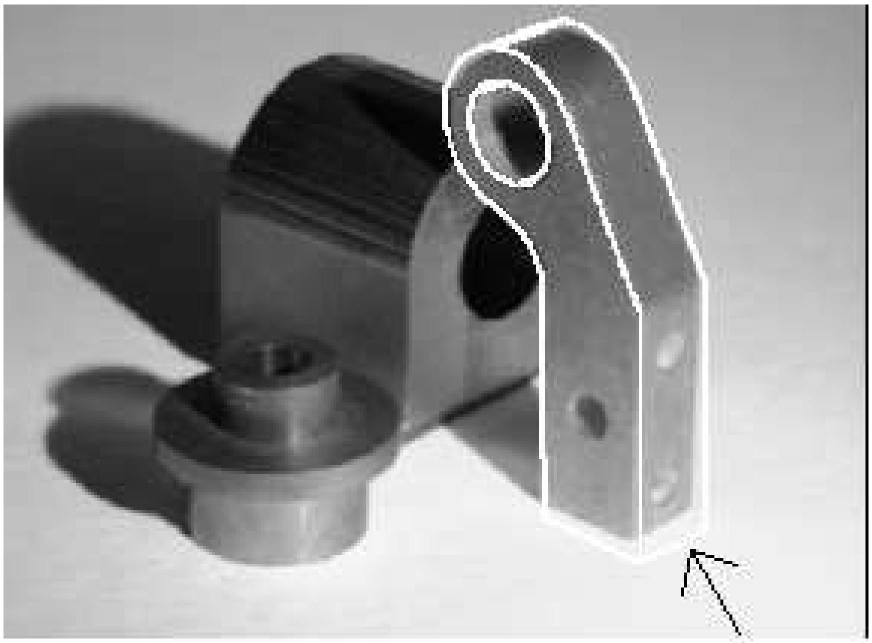

Fig. 15. The alignment between the image of a part and the reprojected outline (white) of using the exterior orientation computed using a single circle. It shows a good alignment in the upper portion of the object where the circle is located and a poor alignment in the lower part (as indicated by the arrow).

Table 2
Iterations needed for the bundle adjustment method to converge using the solutions of the 3 linear methods as initial estimates

Method	Number of iterations
Points only	4
Circle only	6
Points and circle	1

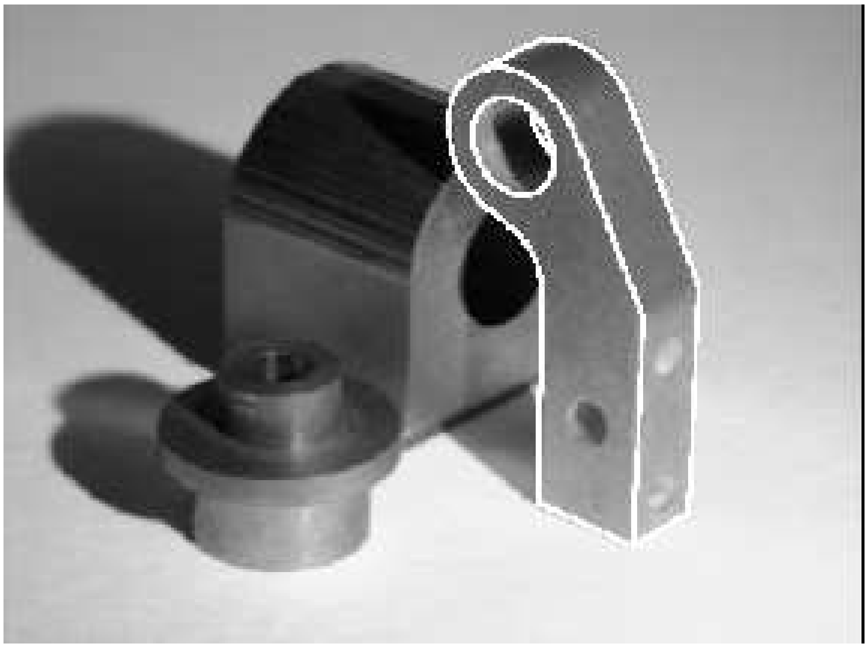

Fig. 16. The result obtained from the bundle adjustment method after only one iteration using the solution in Fig. 13 as the approximation.

in robustness when the number of equations in the linear system reaches the minimum required for a solution to be found. The premise is that one should make use of as many available features as possible in order to improve accuracy and robustness. Our technique follows this principle.

8. Discussion and Summary

In this paper, we presented a linear solution to the exterior orientation estimation problem. The main contributions of this research are the linear framework for fusing information available from different geometric entities and for introducing a novel technique that approximately imposes the orthonormality constraints on the rotation matrix sought. Experimental evaluation using both synthetic data and real images show the effectiveness of our technique for imposing orthonormal constraints in improving estimation errors. The technique is especially effective when the SNR is low or fewer geometric entities are used. The performance study also revealed superiority of the integrated technique to a competing linear technique using only points in terms of robustness and accuracy.

We want to discuss several issues related to the proposed algorithm. First, the algorithm is designed for single images. It can be easily extended to multiple images by cascading the least-squares framework for each image. Second, it is worth to point out that the final accuracy of the estimated exterior orientation depends not only on the measurement accuracy, the feature number, but also heavily on distribution of measurements. In the work described in this paper, we did not address this issue. This will be a future task. Thirdly, the paper emphasizes only the fact that we introduce a framework that is CAPABLE of

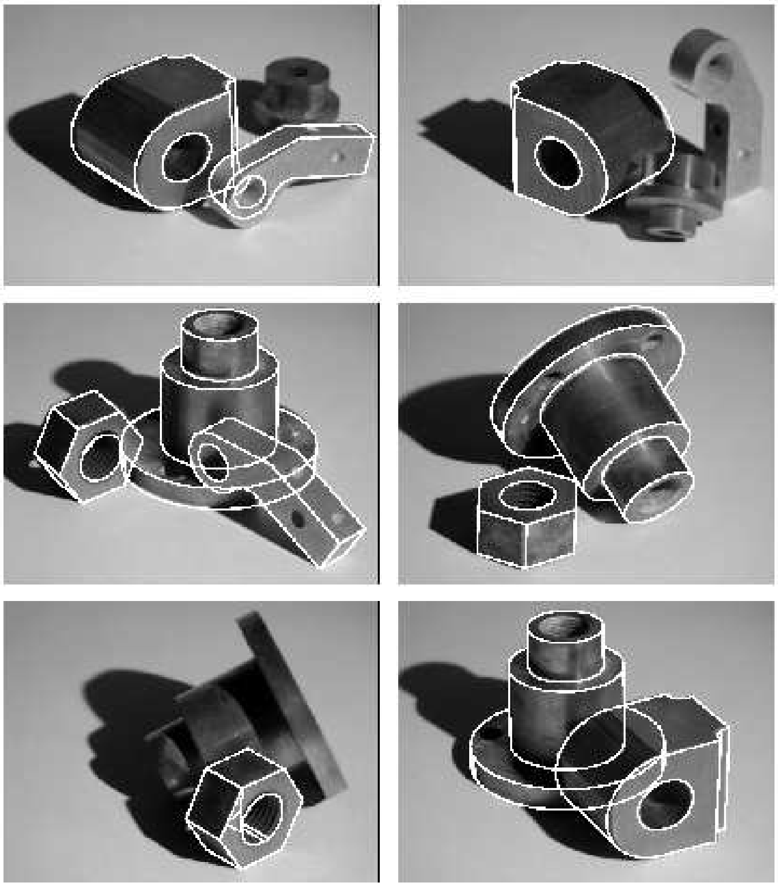

Fig. 17. Results of the integrated method applied to different industrial parts.

integrating different types of geometric features for estimating exterior camera parameters. Studying the effects of what actually cause this improvement is an interesting and nontrivial task. We will leave it for future study. Fourthly, when comparing with point only method with the integrated approach, we used only 10 points. In practice, many more points may be used for the points only method. This can significantly improve the performance of the point only methods and therefore reduce the performance difference between the two. Finally, for the integrated approach, we assume equal weighting for different features. As part of future work, we plan to use error propagation to compute a covariance matrix for each type of feature. The covariance matrices can then used as weightings for different features.

The new technique proposed in this paper is ideal for applications such as industrial automation where robustness, accuracy, computational efficiency, and speed are needed. Its results can also be used as initial estimates in certain applications where more accurate camera parameters are needed. The proposed algorithm is more suitable for images with man-made objects, especially in close-range applications.

References

Abdel-Aziz, Y. I., Karara, H. M., 1971. Direct linear transformation from comparator coordinates into object-space coordinates. ASP Symp. on Close-Range Photogrammetry, 1-18.

Abidi, M. A., 1995. A new efficient and direct solution for pose estimation using quadrangular targets - algorithm and evaluation. IEEE T-PAMI, 17(5), 534-538.

Bopp, H., Krauss, H., 1978. An orientation and calibration method for non-topographic applications. Photogr. Eng. Rem. Sens, 44(9), 1191-1196.

Chen, S.-Y., Tsai, W.-H., 1990. Systematic approach to analytic determination of camera parameters by line features. Pattern Recognition, 23(8), 859-897.

Chen, S. Y., Tsai, W. H., 1991. Determination of robot locations by common object shapes. IEEE Trans. on Robotics and Automation, 7(1), 149-156.

Costa, M. S., 1997. Object recognition and pose estimation using appearance-based features and relational indexing. Ph.D. Dissertation, Intelligent Systems Laboratory, Dept. of Electrical Engineering, University of Washington, Seattle.

Dhome, M. D., Lapreste, J. T., Rives, G., Richetin, M., 1989. Spatial localization of modeled objects of revolution in monocular perspective vision. First European Conference on Computer Vision, 475-485.

Doehler, M., 1975. Verwendung von pass-linien anstelle von pass-punkten in der nahbildmessung. Festschrift K. Schwidefsky, Institute for Photogrammetry and Topography, Univ. of Karlsruhe, Germany, 39-45.

Drewniok, C., Rohr, K., 1997. Exterior orientation - an automatic approach based on fitting analytic landmark models. ISPRS J. of Photogr. and Rem.Sens, 52(3), 132-145.

Echigo, T., 1990. Camera calibration technique using three sets of parallel lines. Machine Vision and Applications, 3(3), 159-167.

Ethrog, U., 1984. Non-metric camera calibration and photo orientation using parallel and perpendicular lines of photographed objects. Photogrammetria, 39(1), 13-22.

Faugeras, O. D., 1993. Three Dimensional Computer Vision: A Geometric Viewpoint. MIT Press.

Finsterwalder, R., 1991. Zur verwendung von passlinien bei photogrammetrischen aufgaben. Zeitschrift fuer Vermessungswesen, 116(2), 60-66.

Fischler, M. A., Bolles, R. C., 1981. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM, 24(6), 381-395.

Forkert, G., 1996. Image orientation exclusively based on free-form tie curves. International Archives of Photogrammetry and Remote Sensing, 31(B3), 196-201.

Forsyth, D., Mundy, J. L., Zisserman, A., Corlho, C., Heller, A., Rothwell, C., 1991. Invariant descriptors for 3d object recognition and pose. IEEE Trans. on Pattern Analysis and Machine Intelligence, 13(10), 971-991.

Gao, M., 1992. Estimating camera parameters from projection of a rectangular parallelepiped. Journal of Northwestern Polytechnic University, 10(4), 427-433.

G.Tommaselli, A. M., Tozzi, C. L., 1996. A recursive approach to space resection using straight lines. Photogr. Eng. Rem. Sens, 62(1), 57-66.

Haralick, R. M., 1988. Determining camera parameters from the perspective projection of rectangle. Pattern Recognition, 22(3), 225-230.

Haralick, R. M., Chu, Y. H., 1984. Solving camera parameters from perspective projection of a parameterized curve. Pattern Recognition, 17(6), 637-645.

Haralick, R. M., Shapiro, L. G., 1993. Computer and Robot Vision, volume 2. Addison-Wesley Publishing Company, Reading, Mass.

Haralick, R. M., Joo, H., Lee, C., Zhang, X., Vaidya, V., Kim, M., 1989. Pose estimation from corresponding point data. IEEE Tran. on Systems, Man, and Cybernetics, 19(6), 1426-1446.

Haralick, R. M., Lee, C., Ottenberg, K., Nolle, M., 1994. Review and analysis of solutions of the three point perspective pose estimation. International Journal of Computer Vision, 13(3), 331-356.

Heikkila, J., 1990. Update calibration of a photogrammetric station. Int'l Arch. of Photogr. and Rem. Sens, 28(5/2), 1234-1241.

Heikkila, J., 1991. Use of linear features in digital photogrammetry. The Photogrammetric Journal of Finland, 12(2), 40-56.

Holt, R. J., Netravali, A. N., 1991. Camera calibration problem: some new results. Computer Vision, Graphics, and Image Processing, 54(3), 368-383.

Horn, B. K. P., 1987. Closed-form solution of absolute orientation using qauternions. J. of Optical Society Am. A., 4(4), 629-642.

Hung, Y., Yeh, P. S., Harwood, D., 1985. Passive ranging to known planar point sets. Proceedings of IEEE Int. Conf. on Robotics and Automation, 80-85.

Ji, Q., Costa, M. S., 1997. New linear techniques for pose estimation using point correspondences. Technical report, Intelligent Systems Lab., #ISL-04-97.

Kager, H., 1989. A universal photogrammetric adjustment system. Optical 3-D Measurement Techniques, 447-455.

Kamgar-Parsi, B., Eas, R. D., 1990. Calibration of stereo system with small relative angles. Computer Vision, Graphics, and Image Processing, 51(1), 1-19.

Kruck, E., 1984. A program for bundle adjustment for engineering applications - possibilities, facilities and practical results. Int'l Arch. of Photogr. and Rem. Sens, 25(A5), 471-480.

Kumar, R., Hanson, A. R., 1989. Robust estimation of camera location and orientation from noisy data having outliers. IEEE Workshop on Interpretation of 3D Scenes, 52-60.

Lee, R., Lu, P. C., Tsai, W. H., 1990. Robot location using single views of rectangular shapes. Photogr. Eng. Rem. Sens, 56(2), 231-238.

Linnainmaa, S., Harwood, D., 1988. Pose determination of a three-dimensional object using triangle pairs. IEEE T-PAMI, 10(11), 634-647.

Liu, Y., Huang, T. S., Faugeras, O. D., 1990. Determination of camera locations from 2d to 3d line and point correspondence. IEEE Trans. on Pattern Analysis and Machine Intelligence, 12(1), 28-37.

Lugnani, J. B., 1980. Using digital entities as control. Ph.D thesis, Dept. of Surveying Eng., The Univ. of New Brunswick, Fredericton.

Ma, D. S., 1993. Conics-based stereo, motion estimation, and pose determination. International Journal of Computer Vision, 10(1), 7-25.

Maas, H. G., 1999. Image sequence based automatic multi-camera system calibration techniques. ISPRS J. of Photogr. and Rem. Sens, 54(6), 352-359.

Masry, S. E., 1981. Digital mapping using entities: a new concept. Photogr. Eng. Rem. Sens, 48(11), 1561-1599.

Mikhail, E. M., 1993. Linear features for photogrammetric restitution and other object completion. Proc. SPIE, 1944, 16-30.

Mikhail, E. M., Mulawa, D. C., 1985. Geometric form fitting in industrial metrology using computer-assisted theodolites. ASP/ACSM Fall Meeting.

Mulawa, D., 1989. Estimation and photogrammetric treatment of linear features. Ph.D. dissertation, School of Civil Eng., Purdue University, West Lafayette, USA.

Mulawa, D. C., Mikhail, E. M., 1988. Photogrammetric treatment of linear features. Int'l Arch. of Photogr. and Rem. Sens, 27(B10), 383-393.

Okamoto, A., 1981. Orientation and construction of models, part i: The orientation problem in close-range photogrammetry. Photogr. Eng. Rem. Sens., 47(11), 1615-1626.

Paderes, F. C., Mikhail, E. M., Foerstner, W., 1984. Rectification of single and multiple frames of satellite scanner imagery using points and edges as control. NASA Symp. on Mathematical Pattern Recognition and Image Analysis, Houston.

Petsa, E., Patias, P., 1994a. Formulation and assessement of straight line based algorithms for digital photogrammetry. Int'l Arch. of Photogr.and Rem. Sens, 30(5), 310-317.

Petsa, E., Patias, P., 1994b. Sensor attitude determination using linear features. Int'l Arch. of Photogr.and Rem. Sens, 30(1), 62-70.

Pettersen, A., 1992. Metrology norway system - an on-line industrial photogrammetric system. Int'l Arch. of Photogr. and Rem. Sens, 29(B5), 43-49.

Phong, T. Q., Horward, R., Yassine, A., Tao, P., 1995. Object pose from 2-d to 3d point and line correspondences. Int. Journal of Computer Vision, 15(3), 225-243.

Rothwell, C. A., Zisserman, A., Marinos, C. I., Forsyth, D., Mundy, J. L., 1992. Relative motion and pose from arbitrary plane curves. Image and Vision Computing, 10(4), 251-262.

Shan, J., 1996. An algorithm for object reconstruction without interior orientation. ISPRS J. of Photogr. and Rem. Sens, 51(6), 299-307.

Strunz, G., 1992. Image orientation and quality assessment in feature based photogrammetry. Robust Computer Vision, 27-40.

Sutherland, I. E., 1974. Three-dimensional data input by tablet. Proceedings of the IEEE, 62(4), 453-461.

Szczepanski, W., 1958. Die loesungsvorschlaege fuer den raeumlichen rueckwaertseinschnitt. Deutsche Geodaetische Kommission, Reihe C, (29).

Tommaselli, A. M. G., Lugnani, J. B., 1988. An alternative mathematical model to the collinearity equation using straight features. Int'l Arch. of Photogr. and Rem. Sens, 27(B3), 765-774.

Tsai, R., 1987. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lens. IEEE Journal of Robotics and Automation, 3(4), 323-344.

Wang, L. L., Tsai, W. H., 1990. Computing camera parameters using vanishing-line information from a rectangular parallepiped. Machine Vision and Applications, 3, 129-141.

Wang, X., Xu, G., 1996. Camera parameters estimation and evaluation in active vision system. Pattern Recognition, 29(3), 439-447.

Weng, J., Huang, T. S., Ahuja, N., 1992. Motion and structure from line correspondences closed-form solution, uniqueness and optimization. IEEE T-PAMI, 14(3), 318-336.