

ROBUST STRUCTURE FROM MOTION USING MULTIPLE CONSTRAINTS ¹

Changming Sun (+)
Andrew K. Forrest ()*

(+) *CSIRO Division of Mathematics and Statistics, Locked Bag 17, North Ryde
NSW 2113, Australia*

(*) *Centre for Robotics, Imperial College, London SW7 2BX, UK*

ABSTRACT

This paper introduces a robust method for obtaining the motion parameters of a camera and 3-D object structure using multiple constraints. These constraints include the epipolar constraint, the constraint from the properties of an essential matrix, the relationship between the essential matrix and the epipoles, and the relationship between the corresponding image points and the essential matrix. Several new constraints have been developed in this work. All these constraints can be described by polynomial equations. Synthetic and real image tests have been carried out, and the results show that the new algorithm has much better performance than the existing methods.

1 INTRODUCTION

Structure from motion algorithms are used in the analysis of image motion caused by relative three-dimensional (3-D) movement between the camera and the imaged objects. These algorithms attempt to recover both the 3-D motion of the camera and the 3-D structures of objects. Linear and nonlinear approaches have been proposed in the literature [1] [2] to the problem of estimating the motion parameters of a rigid body from a set of corresponding points. Longuet-Higgins [1] and Tsai and Huang [3] showed that if the scene contains as many as eight points whose images can be located in each projection and these points can be matched, then the relative orientation of the two projections, the vector of the translation of the camera and the structures of the scene can be computed. The linear approach introduces a set of intermediate parameters, called the "essential matrix", linearly related to the data and uniquely leading to the actual motion parameters in a simple way. Jerian and Jain [4] [5] propose the use of a polynomial system of equations, with the unit quaternions representing rotations, to recover motion and structure

¹This work was performed when the first author was at Centre for Robotics and Automated Systems, Imperial College of Science, Technology and Medicine, London SW7 2BX, UK

under perspective projection. The constraints they used are on the rotation matrix expressed as unit quaternions. They do not use the constraints on the essential matrix. Spetsakis and Aloimonos [2] find the essential matrix by minimizing a quadratic function, and the matrix being able to be decomposed to the rotation matrix and the translation vector. The constraint is nonlinear and very difficult to write down analytically. It is our intention in this paper to use multiple constraints for obtaining robust motion parameters.

Section 2 gives a brief review about the linear algorithm. Section 3 introduces the possible constraints to be used for the refinement of motion parameters. Before concluding, section 4 shows experimental results.

2 THE LINEAR ALGORITHM

Introduce a Cartesian coordinate system with the origin at the pin-hole of the camera and z -axis aligned with the optical axis and pointing outward. Let P be a visible point in the scene, $\mathbf{X} = (X, Y, Z)^T$ and $\mathbf{X}' = (X', Y', Z')^T$ be its 3-D coordinates with respects to the two viewpoints, and $\mathbf{x} = (x, y, 1)^T$, $\mathbf{x}' = (x', y', 1)^T$ be the image points. As the two sets of 3-D coordinates are connected by an arbitrary displacement, we may write

$$\mathbf{X}' = \mathbf{R}(\mathbf{X} - \mathbf{T}) \quad (1)$$

where \mathbf{T} is an unknown translation vector $(T_x, T_y, T_z)^T$ and \mathbf{R} is an unknown rigid rotation matrix which should satisfies the following relationship $\mathbf{R}\mathbf{R}^T = \mathbf{I} = \mathbf{R}^T\mathbf{R}$, $\det(\mathbf{R}) = 1$.

An essential matrix \mathbf{Q} is defined in [1] as intermediate parameter to calculate the rotation matrix and translation vector:

$$\mathbf{Q} = \mathbf{R}\mathbf{S} \quad (2)$$

where \mathbf{S} is a skew-symmetric matrix

$$\mathbf{S} = \begin{pmatrix} 0 & T_z & -T_y \\ -T_z & 0 & T_x \\ T_y & -T_x & 0 \end{pmatrix}$$

It follows from Equations 1–2 that

$$\mathbf{X}'^T \mathbf{Q} \mathbf{X} = 0$$

Dividing the above equation by ZZ' we arrive at the desired relationship between the image coordinates:

$$\mathbf{x}'^T \mathbf{Q} \mathbf{x} = 0 \quad (3)$$

From Equation 2, it can be seen that

$$\mathbf{Q}\mathbf{T} = (\mathbf{R}\mathbf{S})\mathbf{T} = \mathbf{0} \quad (4)$$

The translation parameter \mathbf{T} can be obtained by solving this homogeneous equation. The rotation matrix \mathbf{R} and 3-D object structure can be obtained thereafter as in [1]. For real images, the results of the linear algorithm can hardly be trusted. To obtain better results, more constraints are needed.

3 THE CONSTRAINTS

3.1 Epipoles and the \mathbf{Q} Matrix

The centres of projection of two cameras together with any 3-D point define an epipolar plane, whose intersection with the image plane determines an epipolar line. The epipoles are defined as the intersects of the interocular axis and the image planes. Let the left and the right epipoles be located at $(\pi_1, \pi_2)^T$ and $(\pi'_1, \pi'_2)^T$. If we define two vectors $\mathbf{p} = (\pi_1, \pi_2, 1)^T$ and $\mathbf{p}' = (\pi'_1, \pi'_2, 1)^T$, we can say that \mathbf{T} and \mathbf{p} are similar. Therefore we can have Equation 5 based on Equation 4. Again, the equation for \mathbf{p}' can be obtained as in Equation 6.

$$\mathbf{Q}\mathbf{p} = \mathbf{0} \quad (5)$$

$$\mathbf{p}'^T \mathbf{Q} = \mathbf{0} \quad (6)$$

Divide both sides of the matrix Equation 2, and consider the relationship between the translation vector and the epipoles, we have

$$\mathbf{Q}' = \mathbf{R} \begin{pmatrix} 0 & 1 & -\pi_2 \\ -1 & 0 & \pi_1 \\ \pi_2 & -\pi_1 & 0 \end{pmatrix}$$

where \mathbf{Q}' is a scaled matrix from \mathbf{Q} by $1/T_z$. Multiply the above equation by \mathbf{Q}'^T , we have

$$\mathbf{Q}'^T \mathbf{Q}' = \begin{pmatrix} 1 + \pi_2^2 & -\pi_1 \pi_2 & -\pi_1 \\ -\pi_1 \pi_2 & 1 + \pi_1^2 & -\pi_2 \\ -\pi_1 & -\pi_2 & \pi_1^2 + \pi_2^2 \end{pmatrix} \quad (7)$$

From this symmetric matrix, six polynomial equations (three diagonal and three off-diagonal elements) can be set-up with respect to the elements of \mathbf{Q}' and the two elements of the epipole in the left image. Similarly we can obtain the relationship between the \mathbf{Q} matrix and the epipole in the right image:

$$\mathbf{Q}''^T \mathbf{Q}'' = \begin{pmatrix} 1 + \pi_2'^2 & -\pi_1' \pi_2' & -\pi_1' \\ -\pi_1' \pi_2' & 1 + \pi_1'^2 & -\pi_2' \\ -\pi_1' & -\pi_2' & \pi_1'^2 + \pi_2'^2 \end{pmatrix} \quad (8)$$

where $\mathbf{Q}'' = \frac{1}{T'_z} \mathbf{Q}^T$, and T'_z is the third element of the translation vector from right to left coordinate.

3.2 The Epipolar Constraint

The equation of a straight line of slope m passing through (π_1, π_2) is $(\xi_2 - \pi_2) = m(\xi_1 - \pi_1)$. Similarly, denoting by m' the slope of the corresponding epipolar line in the other image, the equation of the latter is $(\xi'_2 - \pi'_2) = m'(\xi'_1 - \pi'_1)$, where $(\xi_1, \xi_2)^T$ and $(\xi'_1, \xi'_2)^T$ are the points on the epipolar lines. Any point on a certain epipolar line in one image can match any points on the corresponding epipolar line in the other image. Given that all matched points obey Eq. 3, one obtains the following equation by inserting the line equation above into the matrix representation of Equation 3,

$$\begin{pmatrix} \xi'_1 & m'(\xi'_1 - \pi'_1) + \pi'_2 & 1 \end{pmatrix} \mathbf{Q} \begin{pmatrix} \xi'_1 \\ m(\xi_1 - \pi_1) + \pi_2 \\ 1 \end{pmatrix} = 0 \quad (9)$$

for all values of ξ_1 and ξ'_1 . The left hand side of the equation is a second order polynomial in ξ_1 and ξ'_1 and can vanish identically if and only if the coefficient of each term vanishes. This yields four equations. The first of them, arising from the vanishing coefficient of the $\xi_1 \xi'_1$ term, gives the relation

$$m' m q_{22} + m q_{12} + m' q_{21} + q_{11} = 0 \quad (10)$$

which governs the slopes of a pair of epipolar lines. For a pair of image points, $(x, y)^T$ and $(x', y')^T$, the slopes of the epipolar lines (m, m') on the two images are:

$$m = \frac{y - \pi_2}{x - \pi_1}, \quad m' = \frac{y' - \pi'_2}{x' - \pi'_1}$$

Substitute these two equations into Eq. 10, we have:

$$\frac{y - \pi_2}{x - \pi_1} = - \frac{(x' - \pi'_1) q_{11} + (y' - \pi'_2) q_{21}}{(x' - \pi'_1) q_{12} + (y' - \pi'_2) q_{21}} \quad (11)$$

This equation gives the relationship among the matching points, epipoles and the elements of the essential matrix.

3.3 The Properties of the \mathbf{Q} Matrix

Huang and Faugeras [6] have proved that a 3×3 matrix is decomposable if and only if one of its singular values is zero and the other two are equal. Let $\mathbf{Q} = \begin{pmatrix} \mathbf{q}_1^T & \mathbf{q}_2^T & \mathbf{q}_3^T \end{pmatrix}^T$, where $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3$ are the row vectors of \mathbf{Q} . Then, the condition “one of the singular values of \mathbf{Q} is zero” is equivalent to

$$\det(\mathbf{Q}) = 0 \quad (12)$$

and the condition “the two singular values of \mathbf{Q} are equal” is equivalent to

$$\|\mathbf{q}_1 \times \mathbf{q}_2\|^2 + \|\mathbf{q}_2 \times \mathbf{q}_3\|^2 + \|\mathbf{q}_3 \times \mathbf{q}_1\|^2 = \frac{1}{4} \left(\|\mathbf{q}_1\|^2 + \|\mathbf{q}_2\|^2 + \|\mathbf{q}_3\|^2 \right)^2 \quad (13)$$

3.4 Parameter Refinement

Put all these constraints together, we have Equation 14, in which the unknowns are $\pi_1, \pi_2, \pi'_1, \pi'_2$, the elements of \mathbf{Q} matrix. It can be seen from these equations that all of them can be converted to polynomial equations. Therefore, it is easy to obtain the first derivatives for the Newton-Raphson refinement.

$$\left\{ \begin{array}{l} \mathbf{Q}\mathbf{p} = \mathbf{0} \\ \mathbf{p}'^T\mathbf{Q} = \mathbf{0} \\ \det(\mathbf{Q}) = 0 \\ \mathbf{x}'^T\mathbf{Q}\mathbf{x} = 0 \\ \|\mathbf{q}_1 \times \mathbf{q}_2\|^2 + \|\mathbf{q}_2 \times \mathbf{q}_3\|^2 + \|\mathbf{q}_3 \times \mathbf{q}_1\|^2 = \frac{1}{4}(\|\mathbf{q}_1\|^2 + \|\mathbf{q}_2\|^2 + \|\mathbf{q}_3\|^2)^2 \\ \frac{y-\pi_2}{x-\pi_1} = -\frac{(x'-\pi'_1)q_{11}+(y'-\pi'_2)q_{21}}{(x'-\pi'_1)q_{12}+(y'-\pi'_2)q_{21}} \\ \mathbf{Q}^T\mathbf{Q}' = \begin{pmatrix} 1 + \pi_2^2 & -\pi_1\pi_2 & -\pi_1 \\ -\pi_1\pi_2 & 1 + \pi_1^2 & -\pi_2 \\ -\pi_1 & -\pi_2 & \pi_1^2 + \pi_2^2 \end{pmatrix} \\ \mathbf{Q}''^T\mathbf{Q}'' = \begin{pmatrix} 1 + \pi_2'^2 & -\pi_1'\pi_2' & -\pi_1' \\ -\pi_1'\pi_2' & 1 + \pi_1'^2 & -\pi_2' \\ -\pi_1' & -\pi_2' & \pi_1'^2 + \pi_2'^2 \end{pmatrix} \end{array} \right. \quad (14)$$

Assuming that there are N matched points, there will be $2N+20$ equations in Eq. 14. Solving these $2N+20$ nonlinear equations by Newton-Raphson's method, we can get the positions of the epipoles on the two image plane and the improved \mathbf{Q} matrix. The epipolar lines thus obtained are more accurate than those obtained just from Eqs. 5 and 6. In [7] and [8], Braccini used Eqs. 12 and 13 as rigid constraint to obtain more reliable \mathbf{Q} matrix compared to the linear algorithm. By using more constraints as in Equation 14, it is expected that the results can be improved further.

4 EXPERIMENTAL RESULTS

To verify the analysis developed above, simulations have been performed. The number of points used was 30, and the noise variance is 0.0833. The exact values and the results obtained from different (linear, Braccini's and the robust) algorithms about the translation vector \mathbf{T} , the rotation angle \mathbf{R} and the left and right epipoles are given in Table 1. It can be seen from this table that the results of Braccini's algorithm out-performs the linear algorithm, and the robust algorithm gives better results than both the linear algorithm and the Braccini's algorithm.

For real image testing, Figure 1(a) shows the images and the edges obtained. The matching of the features is done using structure matching method [9]. From the corresponding points obtained, the essential matrix is calculated using the linear method described in section 2 and the epipolar lines are obtained from this essential \mathbf{Q} matrix. The corresponding epipolar lines are shown in Figure 1(b)(top two). It can be seen that the epipolar line in the right image is far from the true position.

If we use the improved method (solving $2N+20$ equations at the same time) to calculate the epipolar lines, the result is shown in Figure 1**(b)** (bottom two). The improvement of the result can easily be seen: the resultant epipolar line passes through the true corresponding point. The obtained epipolar lines can be used for matching purpose.

Table 1: *The exact and obtained values of \mathbf{T} , \mathbf{R} and epipoles.*

	\mathbf{T}	\mathbf{R}	Left Epipole	Right Epipole
Exact	(0.965, 0.000, 0.259)	(-0.000, 0.000, -30.000)	(3.732, 0.000)	(-3.732, -0.000)
Linear	(0.939, 0.003, 0.343)	(-0.190, 0.132, -30.350)	(2.738, 0.009)	(-5.647, -0.034)
Braccini	(0.954, 0.003, 0.299)	(-0.171, 0.107, -30.091)	(3.193, 0.009)	(-3.193, -0.009)
Robust	(0.964, 0.003, 0.267)	(-0.121, 0.091, -30.995)	(3.605, 0.011)	(-3.605, -0.017)

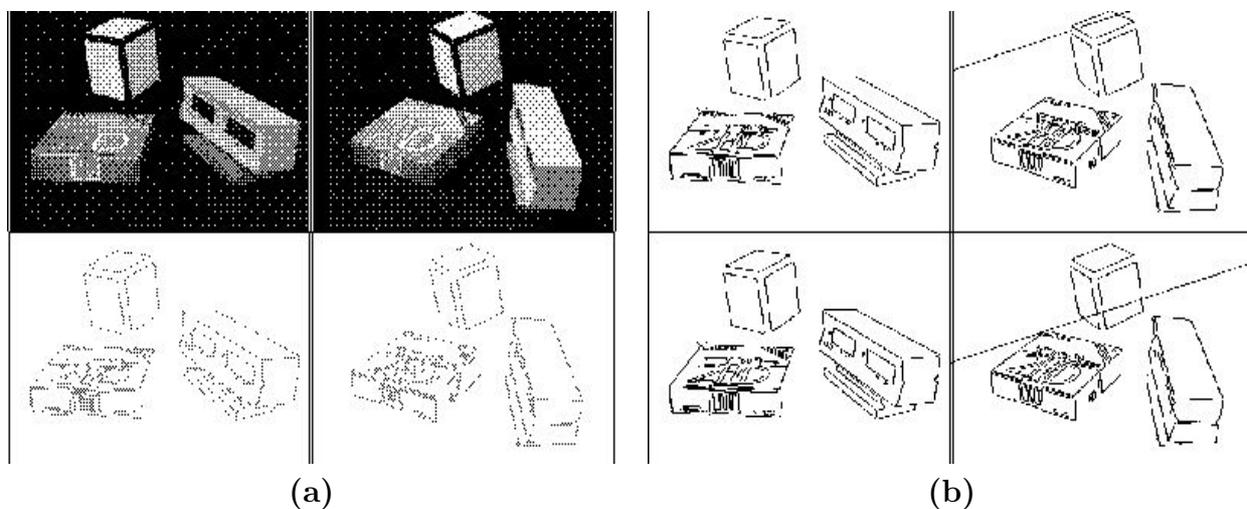


Figure 1: **(a)** *Two images taken by one camera at two positions (top two) and the edges obtained from them (bottom two);* **(b)** *The epipolar lines obtained just from the essential matrix (top two) and the epipolar lines obtained after refinement using multiple constraints (bottom two).*

Figure 2**(a)** shows another set of images and their edges. Figure 2**(b)** illustrates the four projections of the 3-D results obtained by the linear algorithm; while Figure 2**(c)** gives the four projections of the results obtained by the robust algorithm. The objects in the scene are two cubes and a digital clock. The surfaces are mostly squares and rectangles. It can be seen from the top left projection of Figure 2**(b)** that the squares and the rectangles are skewed to the right, and from the top right projection of the same figure that one of the squares on the surface of the cube has been extended to a rectangle when it should be square. The results given in Figure 2**(c)** more resemble to the real objects.

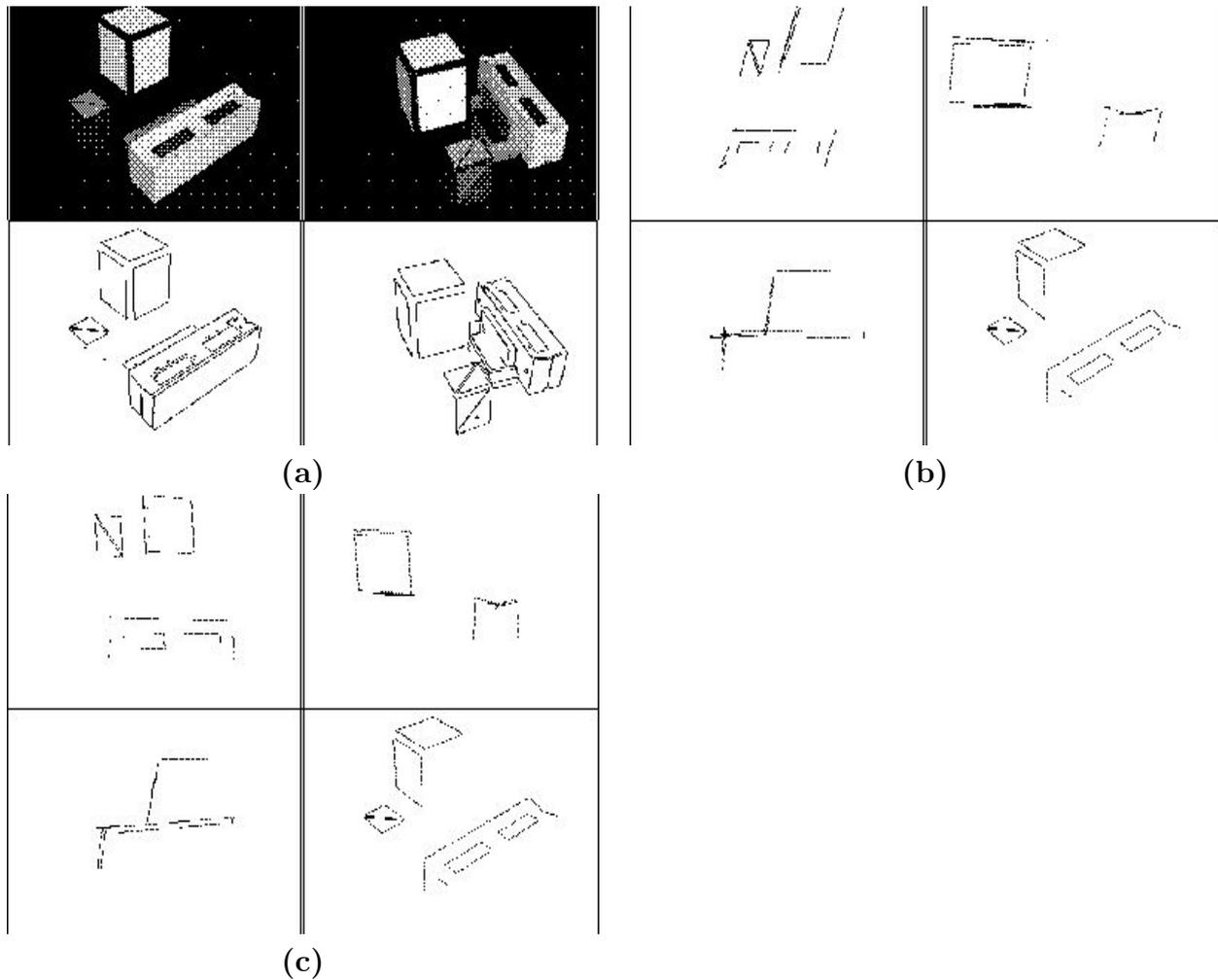


Figure 2: **(a)** Another set of images taken by one camera at two positions (top two) and the edges obtained (bottom two); **(b)** Four projections of the 3-D result obtained from the linear algorithm; **(c)** Four projections of the 3-D result obtained from the robust algorithm.

5 CONCLUSION

Several new constraints have been developed, and they are found to be effective for calculating better motion parameters. These accurate epipoles can be used to obtain the epipolar lines for the ease of matching in a later stage. The improved translation vector is therefore used to compute the rotation matrix, and to calculate the 3-D structures of objects. The results are much better than those obtained just by linear algorithm. In order to obtain better results, more constraints or more information are needed. The constraints about line correspondence might be integrated with the point correspondence approach for the calculation of motion parameters. More constraints usually introduce more computation burden as in a high nonlinearity system, therefore a compromise is frequently made.

6 REFERENCES

- 1) H. C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Nature*, 293:133–135, 1981.
- 2) M. E. Spetsakis and J. Aloimonos. Closed form solution to the structure from motion problem from line correspondences. In *National Conference on AI*, volume 2, pages 738–743, 1987.
- 3) R. Y. Tsai and T. S. Huang. Uniqueness and estimation of three-dimensional motion parameters of rigid objects with curved surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-6:13–26, January 1984.
- 4) C. P. Jerian and R. Jain. Polynomial methods for structure from motion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(12):1150–1166, December 1990.
- 5) C. P. Jerian and R. Jain. Structure from motion—a critical analysis of methods. *IEEE Transactions on Systems, Man and Cybernetics, Part B: Cybernetics*, 21(3):572–588, May/June 1991.
- 6) T. S. Huang and O. D. Faugeras. Some properties of the E matrix in two-view motion estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11:1310–1312, December 1989.
- 7) C. Braccini, G. Gabardella, A. Grattarola, G. Pozzo, and S. Zappatore. Improving the linear approach to motion estimation of rigid bodies by means of nonlinear constraints. In *SPIE Conference on Image Coding*, pages 97–105, 1985.
- 8) C. Braccini, G. Gabardella, A. Grattarola, G. Pozzo, and S. Zappatore. Motion estimation of rigid bodies: Effects of the rigidity constraints. In I. T. Young *et al*, editor, *The Third European Signal Processing Conf., Signal Processing III: Theories and Applications*, pages 645–648. Elsevier Science Publishers B.V. (North-Holland), 1986.
- 9) C. Sun. *A Robust Estimation Method for Three-Dimensional Structure from Multiple Images*. PhD thesis, Centre for Robotics and Automatic Systems, Imperial College London, Exhibition Road, London SW7 2BX, UK, August 1992.