3D-Pose Estimation from a Quadratic-Curved Feature in Two Perspective Views

R. Safaee–Rad *I. Tchoukanov* *B. Benhabib* *K.C. Smith*

Computer Integrated Manufacturing Laboratory, Department of Mechanical Engineering
University of Toronto, 5 King's College Road, Ontario, M5S 1A4, CANADA

ABSTRACT - Estimation of 3D information from 2D-image coordinates is a fundamental problem both in machine vision and computer vision. In this paper, a closed-form analytical solution is presented to the problem of 3D-location estimation of quadratic-curved features based on two images. Compared to previous solution methods, the proposed method has the following advantages: (1) it is a closed-form solution method; (2) it is mathematically simpler; and (3) a priori knowledge of the size and shape of the feature is not required. The method is demonstrated here for an elliptical feature using a simulated experimental set-up.

1. Introduction

Estimation of 3D information from 2D-image coordinates is a fundamental problem in both machine vision and computer vision. The problem of 3D-location (3D-pose) estimation in a scene has been addressed extensively in the applied literature where it is referred to as "feature-based 3D-location estimation of objects". The body of literature dealing with this general problem is concerned with developing mathematical methods for the estimation of an object's location based on **point** features, whether it is a 3-point problem [1], a 4-point problem [2], or an *n*-point problem [3]. Estimation of the object's location based on **line** features, whose mathematics is similar to that of point features, has also been studied [4].

Quadratic-curved features are also very common in machine vision and computer vision, though they are obviously more complex. Quadratic-curve-related situations (involving conic sections) occur in various forms and in various contexts. The fact that quadratic curves are common and important in both the machine-vision and computer-vision fields, is due to several factors: (1) many manufactured objects are bounded by quadratic surfaces or have quadratic-curved edges; (2) quadratic curves or surfaces have been used as artificial landmarks in many machine-vision-related problems; (3) under both orthographic and perspective projections quadratic curves always map onto quadratic curves [5]; (4) with the obvious exception of first-order approximation (straight-line fitting), second-order (segmental or piecewise) approximation of curves is computationally the cheapest and most common approach. Basically, one can classify quadratic-curve-related problems into three kinds, depending on the context in which they occur: (a) pattern recognition and scene analysis, (b) machine-vision metrology, and (c) 3D-location estimation in both the direct and inverse forms.

Haralick and Chu [6] have developed a general method for 3D-location estimation based on quadratic-curved features. They consider two cases: either the shape and the size of a quadratic-curved feature are both known, or only its shape is known. For the first case, their method leads to a highly nonlinear optimization problem -- a nonlinear least-squares problem --, with *three* orientation variables; For the second case, their method leads to a similar nonlinear optimization problem but with *five* variables (three orientation and two size parameters). The main drawback of this method is the residual ambiguity concerning the number of *possible* solutions that exist, and the selection of the one which is *acceptable*. This is due to the nonlinearity of the optimization process which generally yields a *local* minimum depending on the initial estimates. As well, the method does not provide a geometrical interpretation of the problem and its solutions; furthermore, circles, being the most common quadratic feature, are not addressed.

We have previously addressed the *circular*-feature problem and derived a closed-form mathematical solution to the general form of the problem based on 3D-analytical geometry [7-8]. It was shown that based on a *single* image of the feature, there exist two possible solutions to the problem (assuming its radius is known). As well, we have previously addressed the problem of 3D-location estimation based on *quadratic-curved* features [9-10]. There, it was assumed that the true size and the true shape of a feature in object space is known, along with its perspective projection in the image plane (that is, using a *single* image). A set of geometrical constraints on the 3D location for each type of quadratic feature was derived. We proved that, in general, knowledge of the true size and shape does not yield a sufficient number of constraints to *uniquely* determine the 3D orientation and position and that there exist many possible solutions based on a *single* image. We have shown the locus of all possible solutions. As a result, extra constraints must be acquired from other sources and fused with these constraints in order to obtain a unique solution to the 3D-location-estimation problem. Nevertheless, the method developed and the results obtained for quadratic-curved features removed ambiguities that existed in a previously developed solution method [6].

Various types of constraints may be used in order to obtain a unique solution to the 3D-location estimation of quadratic-curved features: a second image of the feature, a coplanar quadratic-curved (QC) feature, a coplanar point feature, a second view along the same optical axis, or a

0-8186-2910-X/92 $3.00 © 1992 IEEE

known distance between the image plane and the QC feature plane. Since acquisition of a second image of the feature is very common in machine vision and computer vision (for example, in stereo vision), in this paper we address the problem with this particular constraint, and derive a corresponding closed-form analytical solution which leads to the *unique* solution to the problem.

2. Analytical Formulation of the Problem

The analytical method proposed in this paper is based on the decomposition of the 3D-location-estimation problem into two parts: first, the 3D orientation of the quadratic-curved-feature's plane is estimated; subsequently, based on the estimated orientation (i.e., feature's surface normal), the 3D position and the third orientation parameter of the feature are calculated.

The problem of 3D-orientation estimation of a quadratic-curved feature based on two images can be stated as follows: Given the effective focal length of a camera, two images of a quadratic-curved feature, and the transformation between the two view points (from information concerning the 3D locations of the camera corresponding to the two views), it is required to estimate the 3D orientation of the feature's plane with respect to the camera frame.

This problem is equivalent to the following one: Given two 3D-cone surfaces defined by two bases (the perspective projections of a quadratic-curved feature in two image planes) and two vertices (the origin of the camera frame corresponding to each of the two views) with respect to a reference frame (for example, the first camera frame), it is required to determine the 3D orientation of a plane (with respect to the same reference frame) which intersects the two cones and generates a *common* curve identical to the quadratic-curved feature.

The equation of the cone whose vertex is the point (α, β, γ) and whose base is defined by a conic

$$\begin{cases} F(x,y) \equiv a'x^2 + 2h'xy + b'y^2 + 2g'x + 2f'y + d' = 0 \\ z = 0\,, \end{cases} \tag{1}$$

can be expressed as [11]:

$$z^2F(\alpha,\beta) - z\gamma\left[x\,\frac{\partial F}{\partial x} + y\,\frac{\partial F}{\partial y} + t\,\frac{\partial F}{\partial t}\right] + \gamma^2 F(x,y) = 0. \tag{2}$$

In equation (2), t is an auxiliary variable by which $F(x,y)$ is made homogeneous $(F(x,y,t))$. The terms $(\partial F/\partial x)$ and $(\partial F/\partial y)$ are evaluated at $x=\alpha$ and $y=\beta$; and the term $(t\,(\partial F/\partial t))$ is calculated by first taking the derivative of the homogeneous equation with respect to t and then equating t to unity and $x=\alpha$ and $y=\beta$. The general form of the equation of a cone that is derived using the above equation is as follows:

$$ax^2 + by^2 + cz^2 + 2fyz + 2gzx + 2hxy + 2ux + 2vy + 2wz + d = 0\,. \tag{3}$$

3. An Analytical Solution of the Problem

The solution method proposed is based on two theorems which were proven in [11-12].

Theorem (1): The curve of intersection of two conicoids is of the fourth degree (a quartic curve). Furthermore, the quartic curve of intersection of two conicoids may consists of: (a) a straight line and a cubic curve, (b) two straight lines and a conic curve, (c) four straight lines, or (d) two conics.

In our case, we know that the two constructed cones have a common conic. Thus, they should have either (1) another common conic; or (2) two other common straight lines (that is, where the second common conic degenerates into two straight lines).

As a result, in our case, the two intersecting cones have *two common planar intersections*. Based on this, we can use the following theorem:

Theorem (2): If $\mathbf{S}_1 = 0$ is the equation of one conicoid, and $\mathbf{P}_1$ and $\mathbf{P}_2$ represent the planes of the common sections, the equation of the other conicoid is of the form:

$$\mathbf{S}_1 + \lambda\,\mathbf{P}_1\,\mathbf{P}_2 = 0\,, \tag{4}$$

where λ is a linear variable coefficient.

Based on this theorem, we can state the following: If $\mathbf{S}_1 = 0$ and $\mathbf{S}_2 = 0$ are the two constructed cones, and $\mathbf{P}_1$ and $\mathbf{P}_2$ are the two common planar sections, then in general the following is true:

$$\mathbf{S}_2 - \lambda_1\,\mathbf{S}_1 = \lambda_2\,\mathbf{P}_1\,\mathbf{P}_2 \tag{5}$$

where,

$$\begin{aligned}
\mathbf{S}_1 &= a_1x^2 + b_1y^2 + c_1z^2 + 2f_1yz + 2g_1zx + 2h_1xy + \\
&\qquad 2u_1x + 2v_1y + 2w_1z + d_1 = 0\,, \\
\mathbf{S}_2 &= a_2x^2 + b_2y^2 + c_2z^2 + 2f_2yz + 2g_2zx + 2h_2xy + \\
&\qquad 2u_2x + 2v_2y + 2w_2z + d_2 = 0\,, \\
\mathbf{P}_1 &= l_1x + m_1y + n_1z + t_1 = 0\,, \\
\mathbf{P}_2 &= l_2x + m_2y + n_2z + t_2 = 0\,.
\end{aligned} \tag{6}$$

Note that, $\mathbf{S}_1$, $\mathbf{S}_2$, $\mathbf{P}_1$, and $\mathbf{P}_2$ are defined with respect to the xyz frame, which is the frame of the first image. From (5) we would obtain ten equations with ten unknowns $(\lambda_1, \lambda_2, l_1, m_1, n_1, t_1, l_2, m_2, n_2, t_2)$.

To simplify the solution of this set of equations, we can proceed, without loss of generality, as follows:

(a) We transfer the first image frame to its *canonical* location (through application of proper 3D-translation and rotation transformations T) [10]. If the canonical frame is defined as XYZ, then the equation of $\mathbf{S}_1$ would reduce to the following:

$$\mathbf{S}_1 = a_1'X^2 + b_1'Y^2 + c_1'Z^2 = 0\,, \tag{7}$$

and for the other equations in (6), we would have (using the transformation T):

$$\mathbf{S}_2 = a_2'X^2 + b_2'Y^2 + c_2'Z^2 + 2f_2'YZ + 2g_2'ZX + 2h_2'XY +$$

$$2u_2'X + 2v_2'Y + 2w_2'Z + d_2' = 0,$$
$$\mathbf{P}_1 = l_1'X + m_1'Y + n_1'Z + t_1' = 0,$$
$$\mathbf{P}_2 = l_2'X + m_2'Y + n_2'Z + t_2' = 0. \quad (8)$$

(b) We normalize $\mathbf{S}_2 - \lambda_1 \mathbf{S}_1$ with respect to d_2' and $\lambda_2 \mathbf{P}_1 \mathbf{P}_2$ with respect to $\lambda_2 t_1' t_2'$ respectively. As a result, the constant terms in the left and right sides of Eq. (5) would be equal to 1.

As a consequence of the above two simplification processes (a) and (b), we obtain from (5) the following set of nine equations with nine unknowns (λ_1, l_1', m_1', n_1', l_2', m_2', n_2', t_1', t_2'), which is simpler to solve:

$$(a_2' - \lambda_1 a_1') / d_2' = (l_1' l_2') / (t_1' t_2')$$
$$(b_2' - \lambda_1 b_1') / d_2' = (m_1' m_2') / (t_1' t_2')$$
$$(c_2' - \lambda_1 c_1') / d_2' = (n_1' n_2') / (t_1' t_2')$$
$$2 f_2' / d_2' = (m_1' n_2' + n_1' m_2') / (t_1' t_2')$$
$$2 g_2' / d_2' = (l_1' n_2' + n_1' l_2') / (t_1' t_2')$$
$$2 h_2' / d_2' = (l_1' m_2' + m_1' l_2') / (t_1' t_2')$$
$$2 u_2' / d_2' = (l_1' t_2' + t_1' l_2') / (t_1' t_2')$$
$$2 v_2' / d_2' = (m_1' t_2' + t_1' m_2') / (t_1' t_2')$$
$$2 w_2' / d_2' = (n_1' t_2' + t_1' n_2') / (t_1' t_2'). \quad (9)$$

To derive a solution set, let

$$a_2'' = a_2' / d_2' \qquad b_2'' = b_2' / d_2' \qquad c_2'' = c_2' / d_2'$$
$$K_5 = 2 f_2' / d_2' \qquad K_6 = 2 g_2' / d_2' \qquad K_7 = 2 h_2' / d_2'$$
$$K_8 = 2 u_2' / d_2' \qquad K_9 = 2 v_2' / d_2' \qquad K_{10} = 2 w_2' / d_2'$$
$$a_1'' = a_1' / d_2' \qquad b_1'' = b_1' / d_2' \qquad c_1'' = c_1' / d_2'$$
$$l_1'' = l_1' / t_1' \qquad m_1'' = m_1' / t_1' \qquad n_1'' = n_1' / t_1'$$
$$l_2'' = l_2' / t_2' \qquad m_2'' = m_2' / t_2' \qquad n_2'' = n_2' / t_2'$$

and

$$K_1 = (b_2'' a_1'' - a_2'' b_1'') / a_1'' \qquad K_2 = b_1'' / a_1''$$
$$K_3 = (c_2'' a_1'' - a_2'' c_1'') / a_1'' \qquad K_4 = c_1'' / a_1''. \quad (10)$$

Then, based on (9) and (10), we can derive the following equation for the unknown l_1'' (11):

$$4 K_2 l_1''^4 - 8 K_2 K_8 l_1''^3 + (-4 K_1 + 5 K_2 K_8^2 + K_9^2) l_1''^2 + (4 K_1 K_8 - K_2 K_8^3 - K_8 K_9^2) l_1'' + (-K_1 K_8^2 - K_7^2 + K_7 K_8 K_9) = 0.$$

To solve equation (11), the method of solution of a quartic equation based on the resolvent cubic equation can be employed [13]. This equation, in general, yields at most four real solutions. However, based on the objective conditions (that is, from Theorem (1), there exist only two common planar sections), it results in only two real solutions, which correspond to the two acceptable solutions for the two planes $\mathbf{P}_1$ and $\mathbf{P}_2$. This will be shown in Section 4 through a numerical example.

Having estimated l_1'', the other coefficients of the two planes $\mathbf{P}_1$ and $\mathbf{P}_2$ can be calculated from the following equations:

$$m_1'' = (K_7 - K_9 l_1'') / (K_8 - 2 l_1'')$$
$$n_1'' = (K_5 - K_{10} m_1'') / (K_9 - 2 m_1'')$$
$$l_2'' = K_8 - l_1'' \qquad m_2'' = K_9 - m_1'' \qquad n_2'' = K_{10} - n_1''. \quad (12)$$

Subsequent to the determination of the six unknowns (l_1'', m_1'', n_1'', l_2'', m_2'', n_2''), we obtain the plane equations of the two common conic sections. Thus, in general, there exist *two* possible solutions to the feature's plane surface normal. However, only one of these solutions is located such that the two vertices of the constructed cones are on *one* side of the plane; and the other solution is located *between* the vertices of the constructed cones. Clearly, the first case is the *acceptable* one, since in a real camera-feature configuration, the feature is always on one side of the camera irrespective of the camera relocation, unless the plane of the feature is *transparent* (that is, the feature can be viewed by the camera from both sides of the feature plane), which is an exceptional case. Thus, the acceptable solution is determined by simply checking whether the intersecting common plane is on the one side of the camera frame in both views, or it is not.

Having determined the equation of the feature's plane, we can determine the feature's 3D position (defined as the feature's center) through the intersection of the first constructed cone with the feature's plane. As well, the third orientation parameter can be estimated (except for a circular feature, since it has universal symmetry with respect to its center).

4. Simulated Experimental Results

Let us consider a situation in which two images of the same elliptical feature are acquired from two different view points. Herein, we *simulate* this situation through the following numerical example.

4.1. Simulated Experimental Set-up

Let the equation of the first viewing cone with respect to the *camera* frame be:

$$12339.385\, x_1^2 + 6300\, y_1^2 + 1296.615\, z_1^2 - 1280.385\, y_1 z_1 - 4898.950\, z_1 x_1 - 11539.230\, x_1 y_1 = 0. \quad (13)$$

Now, if the feature is viewed at its *canonical* location (with respect to the *canonical* camera frame), the equation of the viewing cone would be as follows (using the transformation T between the two frames):

$$16046.344\, X_1^2 + 3942.106\, Y_1^2 - 52.449\, Z_1^2 = 0. \quad (14)$$

Correspondingly, let (13) and (14) for the second view be:

$$136.208\, x_2^2 + 2000\, y_2^2 + 2681.792\, z_2^2 + 3097.056\, y_2 z_2 - 1182\, z_2 x_2 - 297.056\, x_2 y_2 = 0 \quad (15)$$

$$4006.132\, X_2^2 + 836.318\, Y_2^2 - 24.451\, Z_2^2 = 0. \quad (16)$$

The transformation between the canonical camera frames of the second viewing point to the first viewing point is (17):

$$\begin{bmatrix} X_2 \\ Y_2 \\ Z_2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.697 & 0.261 & 0.668 & -22.211 \\ -0.202 & 0.965 & -0.166 & 5.512 \\ -0.688 & -0.019 & 0.726 & -1.128 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \\ 1 \end{bmatrix}.$$

Thus, the equation of the second viewing cone with respect to the first canonical camera frame would be (18):

$$3940.306\,X_1^2 + 2102.481\,Y_1^2 + 3593.207\,Z_1^2 + 2253.008\,Y_1\,Z_1 + 7622.141\,Z_1\,X_1 + 2259.333\,X_1\,Y_1 - 251971.870\,X_1 - 74950.083\,Y_1 - 240660.932\,Z_1 + 4003597.925 = 0.$$

For the above camera-feature configuration, the equation of the elliptical-feature's plane with respect to the first camera frame is:

$$0.250\,x_1 - 0.433\,y_1 - 0.866\,z_1 + 30 = 0 \qquad (19)$$

and with respect to the first canonical camera frame would be:

$$0.324\,X_1 - 0.272\,Y_1 - 0.906\,Z_1 + 30 = 0. \qquad (20)$$

4.2. Experimental Results

In this numerical example, $\mathbf{S}_1$ (7) and $\mathbf{S}_2$ (8) are given by Eqs. (14) and (18). Using (10), we can calculate coefficients K_1 to K_{10} and subsequently solve the fourth order polynomial (11) which yields two real solutions:

$$l_1'' = 0.011 \qquad l_1'' = -0.074\,. \qquad (21)$$

Corresponding to the first solution for l_1'', the other unknowns are estimated as follows (using (12)):

$$m_1'' = -0.009 \quad n_1'' = -0.030 \quad l_2'' = -0.074$$
$$m_2'' = -0.010 \quad n_2'' = -0.030\,. \qquad (22)$$

Similarly, corresponding to the second solution for the l_1'', we obtain the following (using (12)):

$$m_1'' = -0.010 \quad n_1'' = -0.030 \quad l_2'' = 0.011$$
$$m_2'' = -0.009 \quad n_2'' = -0.030\,. \qquad (23)$$

As can be seen, these two sets of solutions are identical, that is, they represent the coefficients of the same two planes $\mathbf{P}_1$ and $\mathbf{P}_2$. This is exactly as expected according to Theorem (1), where there exist only two common intersecting-conic planes. Thus, the equations of these two planes with respect to the first canonical camera frame are:

$$0.011\,X_1 - 0.009\,Y_1 - 0.030\,Z_1 + 1 = 0$$
$$-0.074\,X_1 - 0.010\,Y_1 - 0.030\,Z_1 + 1 = 0 \qquad (24)$$

or

$$\mathbf{P}_1 = 0.324\,X_1 - 0.272\,Y_1 - 0.906\,Z_1 + 30 = 0$$
$$\mathbf{P}_2 = -0.920\,X_1 - 0.120\,Y_1 - 0.373\,Z_1 + 12.474 = 0. \qquad (25)$$

Using the transformation T, we can determine the equations of the both planes with respect to the first camera frame:

$$\mathbf{P}_1 = 0.250\,x_1 - 0.433\,y_1 - 0.866\,z_1 + 30 = 0$$
$$\mathbf{P}_2 = -0.856\,x_1 - 0.511\,y_1 - 0.075\,z_1 + 12.474 = 0. \qquad (26)$$

Since the first plane $\mathbf{P}_1$ is on the one side of the camera frame in two views, it is the acceptable unique solution.

5. Conclusions

A closed-form analytical solution was derived for 3D-location estimation of quadratic-curved features based on two images *without* a priori knowledge about the *size* and *shape* of the feature. It has been shown that there exist, in general, *two* possible solutions to this problem. However, only one of them is acceptable based on the properties of the real camera-feature configuration. A simple method is proposed to determine the acceptable solution. The simulation results support the derived analytical solution method.

References

[1] Fischler, M.A., and Bolles, R.C.., "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography", **Communications of the ACM**, Vol. 24, No. 6, pp. 381-395, June 1981.

[2] Horaud, R., Conio, B., and Leboulleux, O., "An Analytic Solution for the Perspective 4-Point Problem", **Computer Vision, Graphics and Image Processing**, Vol. 47, pp. 33-44, 1989.

[3] Haralick, R.M., and Hyonam, J., "2D-3D Pose Estimation", **IEEE, Proc., 9th Int. Conference on Pattern Recognition**, Vol. 1, pp. 385-391, Nov. 1988.

[4] Ray, L.P., "Estimation of Modeled Object Pose from Monocular Images", **IEEE, Proc., Int. Conference on Robotics and Automation**, Vol. I, pp. 408-413, Cincinnati, Ohio, May 1990.

[5] Hilbert, D., and Cohn-Vossen, S., **Geometry and the Imagination,** First German Edition, 1932, First English Edition, Chelsea Publishing Co., New York, 1952.

[6] Haralick, R.M., "Solving Camera Parameters from the Perspective Projection of a Parameterized Curve", **Pattern Recognition**, Vol. 17, No. 6, pp. 637-645, 1984.

[7] Safaee-Rad, R., Smith, K.C., Benhabib, B., and Tchoukanov, I., "An Analytical Method for the 3D-Location Estimation of Circular Features for an Active-Vision System", **IEEE, Proc., Int. Conference on Systems, Man, and Cybernetics**, pp. 215-220, L.A., CA, Nov. 1990.

[8] Safaee-Rad, R., Smith, K.C., Benhabib, B., and Tchoukanov, I., "3D-Location Estimation of Circular Features for Machine Vision", **IEEE Trans. on Robotics and Automation**, Accepted for Publication, Sept. 1991.

[9] Safaee-Rad, R., Smith, K.C., Benhabib, B., and Tchoukanov, I., "Constraints on Quadratic Curves Under Perspective Projection", **IEEE, Proc., Int. Conference on Systems, Man, and Cybernetics**, Vol. I, pp. 57-62, Charlottesville, Virginia, October 1991.

[10] Safaee-Rad, R., Tchoukanov, I., Smith, K.C., and Benhabib, B., "Constraints on Quadratic-Curved Features Under Perspective Projection", **Image and Vision Computing**, Accepted for Publication, January 1992.

[11] Bell, R.J.T., **An Elementary Treatise on Coordinated Geometry of Three Dimensions**, First Edition, 1910, Third Edition, Macmillan & Co. Ltd., London, 1944.

[12] Sommerville, D.M.Y., **Analytical Geometry of Three Dimensions,** First Edition, Cambridge University Press, London, 1934.

[13] Beyer, W.H., **Standard Mathematical Tables**, CRC Press Inc., Florida, 1981.