JOURNAL OF VISUAL COMMUNICATION AND IMAGE REPRESENTATION
Vol. 5, No. 3, September, pp. 273–281, 1994

How to Catch a Crook

ALFRED M. BRUCKSTEIN,[1] ROBERT J. HOLT, AND ARUN N. NETRAVALI

AT&T Bell Laboratories, Murray Hill, New Jersey 07974

Received February 9, 1993; accepted September 16, 1993

This paper deals with a problem in computer vision: how to recover the motion of a rigid rod, or "crook," thrown in the air towards an observer, from a sequence of images acquired by a pinhole camera. Polynomial equations describing the motion are established, and techniques from algebraic geometry are used to show that in general a sequence of five images is sufficient for the recovery of motion of the rod in accordance with physical laws. Examples are worked out in detail to illustrate our results. © 1994 Academic Press, Inc.

1. INTRODUCTION

Suppose one wants to build a robot that could catch a rigid object thrown towards it. This robot should be able to predict the motion of the object in its free flight, a motion which is governed by well known physical laws. Perhaps the most natural way to provide the necessary data for motion is to equip the robot with one or more "eyes," i.e., electronic cameras providing the image of the object in flight. Anyone who has seen people throwing and catching objects in flight knows that humans can anticipate the future positions of those objects without a great deal of difficulty. Using visual input as well, we seem to have the capability of learning and hard wiring the motion estimation and subsequent object catching movement control process into our brains.

In this paper we consider the problem of recovering the physical motion in space of rigid objects on which some feature points are assumed to be always visible to a pinhole camera imaging system. Once the motion parameters are recovered, the entire trajectory of the object becomes available, enabling a robot system to generate the actions necessary to intercept the free-flying object. An example of a rigid object with permanently visible feature points is a thin rod, or "crook." The endpoints of such "linear" objects are readily available feature points. (See Fig. 1.) We shall show that in the case of an axially symmetric object with two distinguishable feature points on the axis, five image frames at known times are sufficient to recover the motion parameters and predict the future trajectory, assuming noise-free data. More practically, a motion tracking algorithm can be started with as few as five images.

2. SOME PHYSICS OF RIGID BODIES IN FREE FALL

When a rigid body is thrown in the air, its motion is governed by the laws of dynamics. The gravitational force causes the center of mass to move in a conic section, e.g., in the case of uniform gravity a parabola described by

$$\begin{cases} x_{cm}(t) = x_0 + v_x t \\ y_{cm}(t) = y_0 + v_y t \\ z_{cm}(t) = z_0 + v_z t - \dfrac{gt^2}{2}, \end{cases}$$

where (x_0, y_0, z_0) and (v_x, v_y, v_z) are the initial position and velocity, respectively, and g is the acceleration due to gravity. About the center of mass the gravitational forces acting on the rigid body have no torque, hence (under the assumption that air friction etc. are negligible) the motion about the center of mass conserves the angular momentum $\mathbf{L}$. Also, the rotational kinetic energy being constant implies that $\mathbf{L} \cdot \boldsymbol{\omega}(t)$ is constant. Hence $\boldsymbol{\omega}(t)$, the instantaneous angular velocity, having a constant projection on a fixed direction, may wander about in the plane defined by $\mathbf{L} \cdot \boldsymbol{\omega}(t) = \mathbf{L} \cdot \boldsymbol{\omega}(0)$. It turns out to be very difficult to analyze exactly the adventurous motion of $\boldsymbol{\omega}(t)$ in this plane for general bodies characterized by general moments of inertia. However, in the case in which the object has a rotational axis of symmetry one can rather easily deduce that $\boldsymbol{\omega}(t)$ is either constant or has a constant length and precesses about $\mathbf{L}$ on a circle in the plane $\mathbf{L} \cdot \boldsymbol{\omega}(t) =$ constant. (See [10].) This implies that points on the axis of a symmetric object will be moving according to the constraints discussed below.

3. FORMULATION AND NOTATION

The image geometry of the problems we will consider is shown in Fig. 1. A pinhole camera model is used. With-

[1] Permanent address: Dept. of Computer Science, Technion–IIT 32000, Haifa, Israel.

1047-3203/94 $6.00

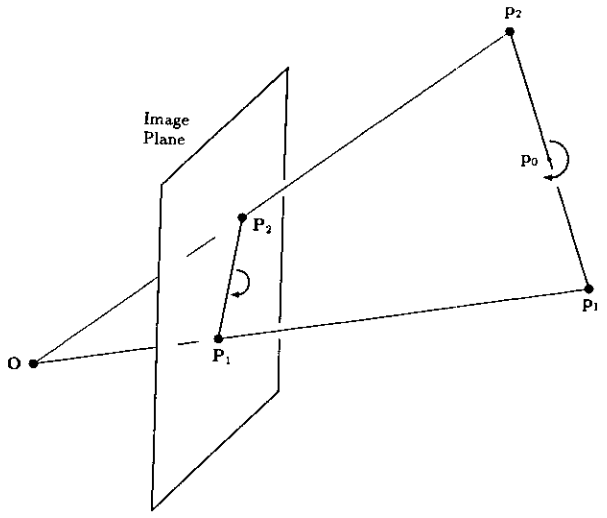

FIG. 1. Perspective imaging geometry for a rotating axially symmetric object. The object is rotating, or precessing, about some line through $\mathbf{p}_0$, the center of mass. $\mathbf{p}_1$ and $\mathbf{p}_2$ are points on the axis of the object, and $\mathbf{P}_1$ and $\mathbf{P}_2$ are their projections onto the image plane.

out loss of generality we can take $z = 1$ to be the image plane. Images are taken of a moving object at evenly spaced times t_j, starting at t_0. It is not necessary for the time intervals to be equal so long as they are all known in order to analyze these problems, but the algebra is considerably simplified with the equal time intervals condition. By processing these images, one attempts to determine both the motion and the structure of the object. The motion parameters are determined by solving equations in which the given data are image coordinates of corresponding points.

Referring to Fig. 1, the following notation will be used:

$\mathbf{p} = (x, y, z)$ = coordinates of a 3-D point at time t_0,
$\mathbf{P} = (X, Y, 1)$ = coordinates of the corresponding image point at t_0,
$\mathbf{p}^{(j)} = (x^{(j)}, y^{(j)}, z^{(j)})$ = coordinates of the same 3-D point at time t_j,
$\mathbf{P}^{(j)} = (X^{(j)}, Y^{(j)}, 1)$ = coordinates of the image point at t_j.

We use the subscript i to denote the ith point, while the superscript (j) denotes j "primes" and refers to points at t_j ($z_i^{(2)} = z_i''$, $z_i^{(3)} = z_i'''$, etc.)

The object and image point coordinates are related by

$$X_i = x_i/z_i, \qquad Y_i = y_i/z_i$$
$$X_i^{(j)} = x_i^{(j)}/z_i^{(j)}, \qquad Y_i^{(j)} = y_i^{(j)}/z_i^{(j)}. \tag{1}$$

4. FIVE VIEWS OF A ROD

In this problem a rod, or any other axially symmetric object with two distinguishable feature points on its axis, is rotating about some axis through its center of mass, while the center of mass itself is moving along some conic section. The situation is illustrated in Fig. 2. Let $\mathbf{C}$ denote a conic section and $\mathbf{C}_j$ the point on the conic at time t_j. Also let the endpoints of the rod be given by $\mathbf{p}_1$ and $\mathbf{p}_2$ initially. Then the center of mass of the rod at time t_j is given not only by $\mathbf{C}_j$ but also by $k\mathbf{p}_1^{(j)} + (1 - k)\mathbf{p}_2^{(j)}$, where $0 < k < 1$. Then the motion of the endpoints is governed by the equations

$$\mathbf{p}_i^{(j)} = \mathbf{C}_j + \mathbf{R}^j(\mathbf{p}_i - \mathbf{C}_0), \qquad i = 1, 2, \tag{2}$$

with the additional constraint that

$$k[\mathbf{p}_1^{(j)} - \mathbf{C}_j] = -(1 - k)[\mathbf{p}_2^{(j)} - \mathbf{C}_j], \tag{3}$$

for some 3×3 rotation matrix $\mathbf{R}$. We will also use

$$\mathbf{q}^{(j)} \triangleq \mathbf{p}_1^{(j)} - \mathbf{p}_2^{(j)} = \mathbf{R}^j(\mathbf{p}_1 - \mathbf{p}_2) = \mathbf{R}^j\mathbf{q}. \tag{4}$$

It is also helpful to consider vectors which are normal to the planes through the $\mathbf{q}^{(j)}$ and the origin. We shall let $\mathbf{N}^{(j)}$ denote a vector perpendicular to the plane through $\mathbf{q}^{(j)}$

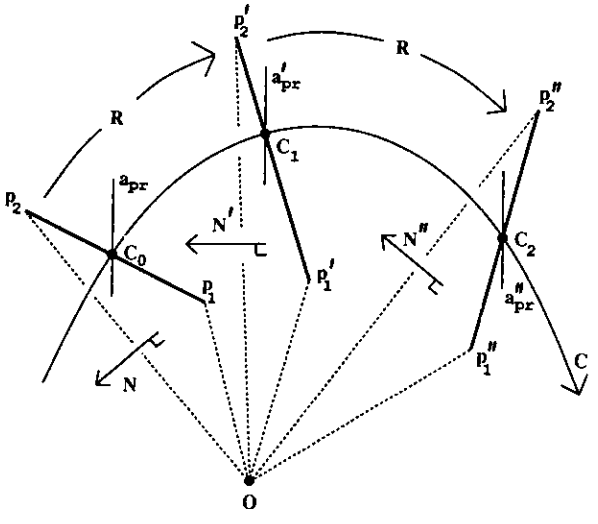

FIG. 2. Three-dimensional motion of a rotating rod. The rod is shown at three time instants by the bold line segments, and its center of mass travels along $\mathbf{C}$, a conic curve. The thin solid lines marked by $\mathbf{a}_{pr}$ represent the axis of precession at the different times; in the absence of external torques this axis has a constant orientation and makes a constant angle with the rod, intersecting the rod at $\mathbf{C}_j$ at time t_j. Each $\mathbf{N}^{(j)}$ is the normal vector to the plane formed by the origin and the rod at time t_j.

and **0**. In terms of the image point coordinates, we define

$$\mathbf{N}^{(j)} = \begin{bmatrix} Y_1^{(j)} - Y_2^{(j)} \\ X_2^{(j)} - X_1^{(j)} \\ X_1^{(j)} Y_2^{(j)} - X_2^{(j)} Y_1^{(j)} \end{bmatrix}. \tag{5}$$

In this problem the observed quantities are the projections of the endpoints onto the image plane $z = 1$. The unknown quantities we wish to determine are the z-components of these feature points, the rotation **R**, and the center of mass of the rod. Knowledge of the center of mass at five time instants allows the determination of the conic section on which it travels. In [4, 13], polynomial equations in the z-components of the feature points were established and solved. The same approach will be used here, along with equations involving the unknown k from Eq. (3) and the components of **R**. It turns out to be useful to express **R** in quaternion form as in [6]. If the matrix **R** is represented by the unit quaternion $\mathring{q} = q_w + iq_x + jq_y + kq_z$, then the components of **R** can be expressed as

$$\mathbf{R} = \begin{bmatrix} q_w^2 + q_x^2 - q_y^2 - q_z^2 & 2(q_xq_y - q_wq_z) & 2(q_xq_z + q_wq_y) \\ 2(q_xq_y + q_wq_z) & q_w^2 - q_x^2 + q_y^2 - q_z^2 & 2(q_yq_z - q_wq_x) \\ 2(q_xq_z - q_wq_y) & 2(q_yq_z + q_wq_x) & q_w^2 - q_x^2 - q_y^2 + q_z^2 \end{bmatrix}, \tag{6}$$

where

$$q_w^2 + q_x^2 + q_y^2 + q_z^2 = 1. \tag{7}$$

In this representation $(q_w, q_x, q_y, q_z) = (\cos(\theta/2), n_x \sin(\theta/2), n_y \sin(\theta/2), n_z \sin(\theta/2))$, where θ is the angle of rotation and $\mathbf{n} = [n_x \quad n_y \quad n_z]^T$ is the axis of rotation of **R**. From (6) it can be directly observed that $\mathring{q}$ and $-\mathring{q}$ represent the same rotation.

We will now establish several polynomial equations relating the unknowns. Useful equations for the determination of **R** can be obtained by making the important observation as in [7–9] that the vectors $\mathbf{RN}^{(j-1)}$, $\mathbf{N}^{(j)}$, and $\mathbf{R}^{-1}\mathbf{N}^{(j+1)}$ are coplanar. Referring to Fig. 2 and Eq. (4), we find that the vectors **Rq**, **q**′, and $\mathbf{R}^{-1}\mathbf{q}''$ are parallel. The vectors **RN**, **N**′, and $\mathbf{R}^{-1}\mathbf{N}''$ are perpendicular to these parallel vectors and consequently are coplanar. Thus the scalar triple product of **RN**, **N**′, and $\mathbf{R}^{-1}\mathbf{N}''$ vanishes, and we have

$$\mathbf{RN}^{(j-1)} \cdot [\mathbf{N}^{(j)} \times \mathbf{R}^{-1}\mathbf{N}^{(j+1)}] = 0, \qquad j = 1, 2, 3. \tag{8}$$

Equation (8) provides three homogeneous quartic polynomials in the variables q_w, q_x, q_y, q_z when $\mathbf{R}^{-1}$ is written as $\mathbf{R}^T$, which is the case for real rotation matrices. From Bézout's theorem one might expect there to be $4^3 = 64$ solutions to this system in $\mathbf{P}^3$, projective 3-space. $\mathbf{P}^3$ is the set of 4-tuples (z_0, z_1, z_2, z_3) besides $(0, 0, 0, 0)$ where points lying on the same line through the origin are identified. However, it was discovered that there were always infinitely many solutions with

$$q_w = 0 \qquad \text{and} \qquad q_x^2 + q_y^2 + q_z^2 = 0. \tag{9}$$

This happens because the polynomials (8) are of the form $f_4q_w^4 + f_3q_w^3 + f_2q_w^2 + f_1q_w + (q_x^2 + q_y^2 + q_z^2)f_0$, where the f_m are polynomials in q_x, q_y, and q_z. These solutions certainly do not satisfy (7), but in order to use the algebraic geometry theorems in the Appendix, which give results concerning the number of solutions, we must have a set of homogeneous polynomials.

This difficulty can be overcome by considering the resultants obtained by eliminating q_x, q_y, and q_z among the polynomials of (8), and then factoring out powers of q_w. Good discussions of resultants can be found in [2, 15]. Referring to Eq. (8), let

$$p_1 = p_1(q_w, q_x, q_y, q_z) = \mathbf{RN} \cdot [\mathbf{N}' \times \mathbf{R}^{-1}\mathbf{N}'']$$
$$p_2 = p_2(q_w, q_x, q_y, q_z) = \mathbf{RN}' \cdot [\mathbf{N}'' \times \mathbf{R}^{-1}\mathbf{N}''']$$
$$p_3 = p_3(q_w, q_x, q_y, q_z) = \mathbf{RN}'' \cdot [\mathbf{N}''' \times \mathbf{R}^{-1}\mathbf{N}^{(4)}].$$

The resultant of two fourth degree polynomials in q_z, say $p_1 = a_4q_z^4 + a_3q_z^3 + a_2q_z^2 + a_1q_z + a_0$ and $p_2 = b_4q_z^4 + b_3q_z^3 + b_2q_z^2 + b_1q_z + b_0$, is the following 8×8 determinant:

$$R(p_1, p_2) = \begin{vmatrix} a_4 & a_3 & a_2 & a_1 & a_0 & 0 & 0 & 0 \\ 0 & a_4 & a_3 & a_2 & a_1 & a_0 & 0 & 0 \\ 0 & 0 & a_4 & a_3 & a_2 & a_1 & a_0 & 0 \\ 0 & 0 & 0 & a_4 & a_3 & a_2 & a_1 & a_0 \\ b_4 & b_3 & b_2 & b_1 & b_0 & 0 & 0 & 0 \\ 0 & b_4 & b_3 & b_2 & b_1 & b_0 & 0 & 0 \\ 0 & 0 & b_4 & b_3 & b_2 & b_1 & b_0 & 0 \\ 0 & 0 & 0 & b_4 & b_3 & b_2 & b_1 & b_0 \end{vmatrix}. \tag{10}$$

Resultants have the important property that $R(p_1, p_2) = 0$ whenever both $p_1 = 0$ and $p_2 = 0$.

In the case at hand, if we eliminate q_z between p_1 and p_2, between p_1 and p_3, and between p_2 and p_3, we obtain three homogeneous 16th degree polynomials in q_w, q_x, and q_y, each of which contains the factor q_w^2. Let the three 14th degree polynomials obtained by removing the factors of q_w^2 be p_{12}, p_{13}, and p_{23}, respectively. All real solutions of (8) must satisfy $p_{12}(w, x, y) = p_{13}(w, x, y) = p_{23}(w, x, y) = 0$. If we now eliminate q_y between p_{12} and p_{13}, between p_{12} and p_{23}, and between p_{13} and p_{23}, and take the greatest common factor of these, we obtain a homogeneous 45th degree polynomial in q_w and q_x which contains a factor of q_w^3. Let the 42nd degree polynomial obtained by removing the factor of q_w^3 be $p^{(42)}(w, x)$. Then real solutions of (8) must satisfy $p^{(42)}(w, x) = 0$. One might worry that a real solution in which q_w does equal zero (a solution with a 180° rotation) would be lost by removing these factors of q_w but this is not the case. Instead, if the data is such that there is a real solution with $q_w = 0$, then $p^{(42)}(w, x)$ is divisible by q_w. Thus in all cases all the real solutions satisfy $p^{(42)}(w, x) = 0$. We can construct polynomials $p^{(42)}(w, y)$ and $p^{(42)}(w, z)$ in an analogous manner.

Hence we will seek solutions of this system:

$$\begin{aligned} p_1(q_w, q_x, q_y, q_z) &= 0 \qquad p_2(q_w, q_x, q_y, q_z) = 0 \qquad p_3(q_w, q_x, q_y, q_z) = 0 \\ p^{(42)}(w, x) &= 0 \qquad p^{(42)}(w, y) = 0 \qquad p^{(42)}(w, z) = 0. \end{aligned} \tag{11}$$

The first three equations of (11) are homogeneous quartics and the last three equations of (11) are homogeneous 42nd degree polynomials in q_w, q_x, q_y, q_z. The above resultant computations were made using the symbolic manipulation program Maple [3].

Solutions for several specific examples of the homogeneous system (11) were found through the use of the algebraic geometry computing package Macaulay [14] and Maple, while the homotopy continuation program Polsys in the package Hompack [12] and the symbolic manipulation program Macsyma [1] were used to verify the results. In each case it was found that there were 42 solutions for (q_w, q_x, q_y, q_z) in $\mathbf{P}^3$. While the algebraic geometry results in the Appendix can be used to show that there are in general 42 solutions to the system (11), we will instead consider more equations in order to reduce the size of the solution space.

In view of Eq. (4) we have $\mathbf{Rq}^{(j)} = \mathbf{q}^{(j+1)}$ for $j = 0, 1, 2, 3$. This provides four vector equations, or equivalently twelve scalar equations:

$$\begin{aligned} &(q_w^2 + q_x^2 - q_y^2 - q_z^2)[X_1^{(j)}z_1^{(j)} - X_2^{(j)}z_2^{(j)}] \\ &\quad + 2(q_xq_y - q_wq_z)[Y_1^{(j)}z_1^{(j)} - Y_2^{(j)}z_2^{(j)}] \\ &\quad + 2(q_xq_z + q_wq_y)[z_1^{(j)} - z_2^{(j)}] = X_1^{(j+1)}z_1^{(j+1)} - X_2^{(j+1)}z_2^{(j+1)} \\ &2(q_xq_y + q_wq_z)[X_1^{(j)}z_1^{(j)} - X_2^{(j)}z_2^{(j)}] \\ &\quad + (q_w^2 - q_x^2 + q_y^2 - q_z^2)[Y_1^{(j)}z_1^{(j)} - Y_2^{(j)}z_2^{(j)}] \\ &\quad + 2(q_yq_z - q_wq_x)[z_1^{(j)} - z_2^{(j)}] = Y_1^{(j+1)}z_1^{(j+1)} - Y_2^{(j+1)}z_2^{(j+1)} \\ &2(q_xq_z - q_wq_y)[X_1^{(j)}z_1^{(j)} - X_2^{(j)}z_2^{(j)}] \\ &\quad + 2(q_yq_z + q_wq_x)[Y_1^{(j)}z_1^{(j)} - Y_2^{(j)}z_2^{(j)}] \\ &\quad + (q_w^2 - q_x^2 - q_y^2 + q_z^2)[z_1^{(j)} - z_2^{(j)}] = z_1^{(j+1)} - z_2^{(j+1)}, \end{aligned} \tag{12}$$

$$j = 0, 1, 2, 3.$$

We still need equations that determine the center of mass of the rod. Recall that the center of mass at time t_j is given by $k\mathbf{p}_1^{(j)} + (1 - k)\mathbf{p}_2^{(j)}$. Then equations that constrain the center of mass to travel within a single plane take the form

$$\text{Rank} \begin{bmatrix} kX_1z_1 + (1-k)X_2z_2 & kY_1z_1 + (1-k)Y_2z_2 & kz_1 + (1-k)z_2 & 1 \\ kX_1'z_1' + (1-k)X_2'z_2' & kY_1'z_1' + (1-k)Y_2'z_2' & kz_1' + (1-k)z_2' & 1 \\ kX_1''z_1'' + (1-k)X_2''z_2'' & kY_1''z_1'' + (1-k)Y_2''z_2'' & kz_1'' + (1-k)z_2'' & 1 \\ kX_1'''z_1''' + (1-k)X_2'''z_2''' & kY_1'''z_1''' + (1-k)Y_2'''z_2''' & kz_1''' + (1-k)z_2''' & 1 \\ kX_1^{(4)}z_1^{(4)} + (1-k)X_2^{(4)}z_2^{(4)} & kY_1^{(4)}z_1^{(4)} + (1-k)Y_2^{(4)}z_2^{(4)} & kz_1^{(4)} + (1-k)z_2^{(4)} & 1 \end{bmatrix} = 3, \tag{13}$$

or in other words, each of the five 4×4 subdeterminants of (13) is zero. However, considering the coefficients of k^3 in the subdeterminants of (13), we observe that these coefficients are zero exactly when the tips of the vectors $\mathbf{q}$, $\mathbf{q}'$, $\mathbf{q}''$, $\mathbf{q}'''$, $\mathbf{q}^{(4)}$ are coplanar, which is the case since $\mathbf{Rq}^{(j)} = \mathbf{q}^{(j+1)}$, as mentioned above. Therefore we can

drop the coefficients of k^3 in (13) and consider the subdeterminants to be quadratic, not cubic, equations in k.

As stated above, in order to use the algebraic geometry theorems in the Appendix we must have a system of homogeneous equations. To this end we introduce variables q_0 and k_0. Then the homogeneous versions of Eqs. (7), (12), and (13) are

$$q_w^2 + q_x^2 + q_y^2 + q_z^2 = q_0^2, \tag{14}$$

$$\mathbf{R}\mathbf{q}^{(j)} = \mathbf{q}^{(j+1)}q_0^2, \tag{15}$$

and

$$\text{Rank}\begin{bmatrix} k\mathbf{p}_1 + (k_0 - k)\mathbf{p}_2 & 1 \\ k\mathbf{p}_1' + (k_0 - k)\mathbf{p}_2' & 1 \\ k\mathbf{p}_1'' + (k_0 - k)\mathbf{p}_2'' & 1 \\ k\mathbf{p}_1''' + (k_0 - k)\mathbf{p}_2''' & 1 \\ k\mathbf{p}_1^{(4)} + (k_0 - k)\mathbf{p}_2^{(4)} & 1 \end{bmatrix} = 3, \tag{16}$$

respectively. As mentioned above, (16) is to be regarded as a system of five homogeneous quadratics in k and k_0, one equation for each 4×4 subdeterminant.

Solutions for several specific examples of systems consisting of Eqs. (11) and (14)–(16) were found using Macaulay, Maple, Macsyma, and Polsys. In each case it was found that there was a unique solution for the rotation $\mathbf{R}$ and for k, and for all the z-components of the feature points up to a common scale factor. Thus we obtain the position of the center of mass (up to the same scale factor) at each of the five time instants, and that is sufficient to determine the conic on which it travels.

THEOREM 1. *Suppose we are given the perspective projections of a rod, in torque-free rigid motion about its unknown center of mass, at five evenly spaced time intervals. Suppose further that the path of the center of mass is known to be a conic section. Then there is in general a unique solution for the position and rotation of the rod.*

The complete proof is quite involved and can be found in the Appendix.

6. EXAMPLES

EXAMPLE 1. This is an example where five snapshots of a rotating rod are taken at evenly spaced time intervals. The motion of the rod is depicted in Fig. 3, which was generated by Maple. We assume that the center of mass of the rod travels along a conic section, but that the center of mass is not necessarily halfway between the two endpoints. The rotation of the rod is uniquely determined, and its center of mass and the path of the center of mass are uniquely determined up to a common scale factor.

j	0	1	2	3	4
$X_1^{(j)}$	−1/10	1/2	29/40	137/160	371/730
$Y_1^{(j)}$	1/2	7/8	17/40	7/320	53/730
$X_2^{(j)}$	−11/74	−1/10	31/140	163/440	529/1370
$Y_2^{(j)}$	55/74	29/40	43/140	−307/880	−953/1370

With these observations the z-components of the feature points are found to have the following unique (up to a scale factor) solution:

j	0	1	2	3	4
$z_1^{(j)}$	2/3	4/3	8/3	32/15	146/75
$z_2^{(j)}$	74/45	20/9	28/9	88/45	274/225

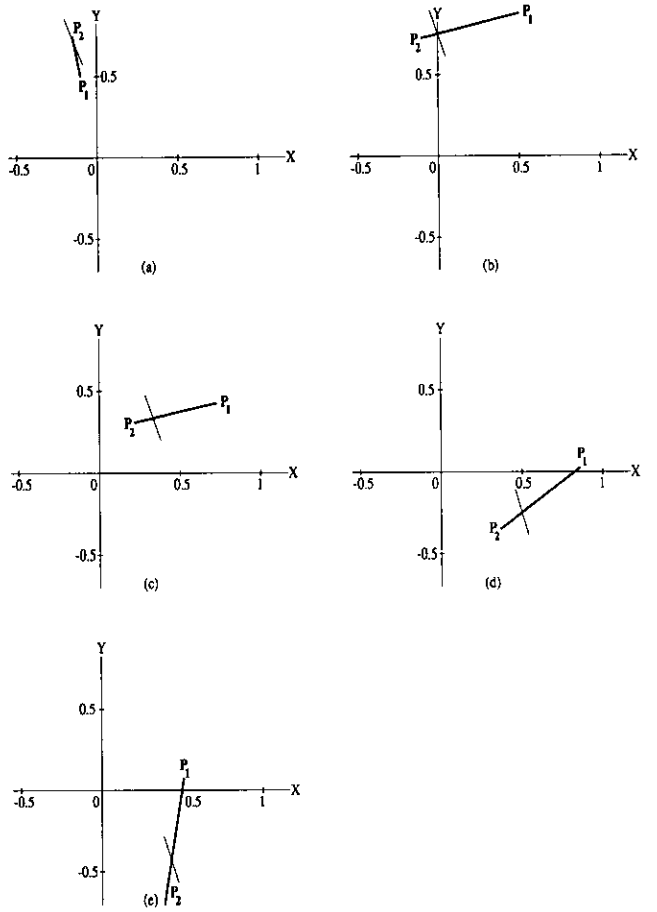

FIG. 3. Image plane observations for Example 1. Parts (a) through (e) are the views at times t_0 through t_4. The bold line segments represent the rotating rod and the light segments represent the axis of rotation. In each frame, the axis of rotation is translated to show it intersecting the center of mass of the rod.

The center of mass is found to lie three-quarters of the way from $\mathbf{p}_1$ to $\mathbf{p}_2$, or at $(\mathbf{p}_1 + 3\mathbf{p}_2)/4$, and travels in the plane $4x + 2y - 3z = -3$ and on the ellipse which is the intersection of this plane with the cylinder $2x^2 + 2xy + y^2 - 4x - 3y + 2 = 0$. Between successive pairs of snapshots the rod is subjected to the rotation

$$\mathbf{R} = \begin{bmatrix} 37/45 & -4/9 & -16/45 \\ 16/45 & 8/9 & -13/45 \\ 4/9 & 1/9 & 8/9 \end{bmatrix}.$$

In this case the axis of rotation, $[1/3 \quad -2/3 \quad 2/3]^T$, is perpendicular to each of the $\mathbf{q}^{(j)} = \mathbf{p}_1^{(j)} - \mathbf{p}_2^{(j)}$, so the rod is rotating end-over-end.

In Fig. 3, the perspective projections account for the differences in the lengths of the images of the rod and the apparent locations of the center of mass, though the center of mass is always about three times as far from $\mathbf{p}_1^{(j)}$ as it is from $\mathbf{p}_2^{(j)}$. The projection of the center of mass onto the image plane $z = 1$ also lies on a conic, in this case the ellipse $13X^2 + 10XY + 4Y^2 - 6X - 3Y = 0$.

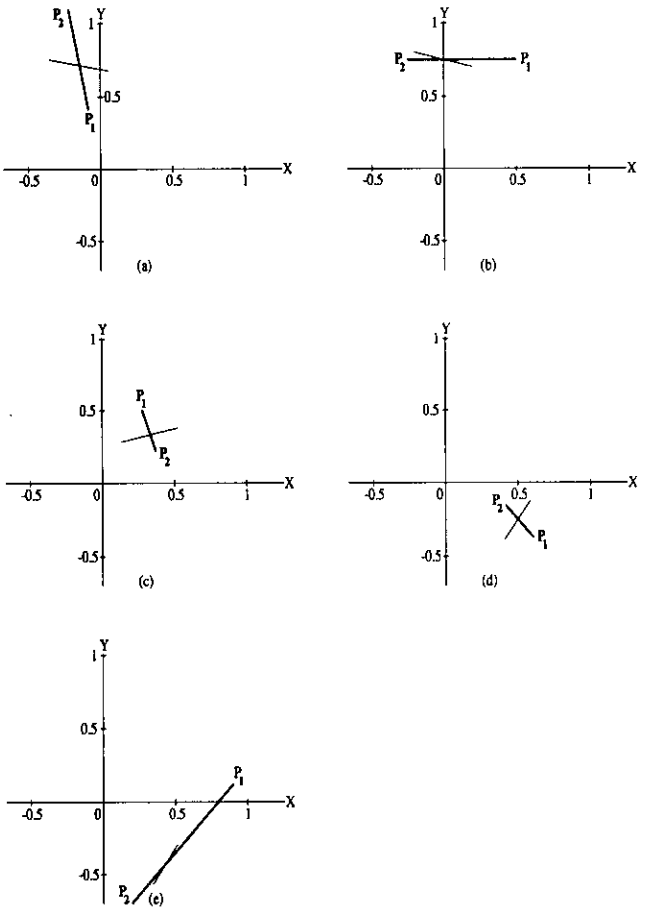

FIG. 4. Image plane observations for Example 2. Parts (a) through (e) are the views at times t_0 through t_4. The bold line segments represent the rotating rod and the light segments represent the axis of rotation, or precession. In each frame, the axis of precession is translated to show it intersecting the center of mass of the rod.

EXAMPLE 2. In this example, the rod is rotating, or precessing, about a line through its center of mass, but the axis of precession is not perpendicular to the rod. The motion of the rod is depicted in Fig. 4, which was also generated by Maple. Once again, the rotation of the rod is uniquely determined, and its center of mass and the path of the center of mass are uniquely determined up to a common scalar multiple.

j	0	1	2	3	4
$X_1^{(j)}$	−1/12	1/2	5/18	20/33	155/172
$Y_1^{(j)}$	5/12	3/4	1/2	−49/132	21/172
$X_2^{(j)}$	−2/9	−1/4	10/27	5/12	70/353
$Y_2^{(j)}$	10/9	3/4	2/9	−13/84	−246/353

With these observations the z-components of the feature points are found to have the following unique (up to a scale factor) solution:

j	0	1	2	3	4
$z_1^{(j)}$	12/5	2	18/5	66/25	172/125
$z_2^{(j)}$	9/10	2	27/10	42/25	353/250

The center of mass is found to lie two-thirds of the way from $\mathbf{p}_1$ to $\mathbf{p}_2$, or at $(\mathbf{p}_1 + 2\mathbf{p}_2)/3$, and travels along the same conic as in Example 1, the intersection of the plane $4x + 2y - 3z = -3$ with the cylinder $2x^2 + 2xy + y^2 - 4x - 3y = -2$. Between successive pairs of snapshots the rod is subjected to the rotation

$$\mathbf{R} = \begin{bmatrix} 0 & 0 & 1 \\ 4/5 & -3/5 & 0 \\ 3/5 & 4/5 & 0 \end{bmatrix}.$$

In this case the axis of precession, $[2/3 \quad 1/3 \quad 2/3]$, makes an angle of $\cos^{-1}(2/3) = 48.2°$ with each of the $\mathbf{q}^{(j)} = \mathbf{p}_1^{(j)} - \mathbf{p}_2^{(j)}$, so the rod is not rotating end-over-end.

In Fig. 4, the apparent different directions of the axis of precession are due to the perspective projection and the proximity of the rod to the camera plane. Also, the rod is nearly parallel to the camera plane at time t_4 and close to perpendicular to that plane at times t_2 and t_3, which accounts for the large difference in the length of the image in Fig. 4e compared with those in Figs. 4c and 4d.

7. CONCLUSIONS

We have shown that the trajectory of a rotating axially symmetric object can be determined provided that two

points on the axis are visible in five snapshots taken by a pinhole camera. This is true generically, i.e., when the conditions are not zero probability events of nongeneral position. In the case of a thin rigid rod, the endpoints are clearly visible and are the natural candidates for the feature points. A crook is not quite axially symmetric, hence the analysis of its motion is more complex, and a crook will indeed be much more difficult to catch. This is very much like real life! We could, however, assume that the asymmetry is not too significant and try to extend to an almost symmetric object the results we have on thin rigid rods.

Note that the results discussed above, although mathematically rigorous and pleasing, should only be regarded as proofs of the fact that in a short sequence of images of the object to be tracked in flight there is indeed enough information to extract the motion parameters and hence exactly predict the entire flight path. Numerical solutions for many robot vision problems based on a minimal amount of information tend to be quite sensitive to noise, and as this is the case here this paper is primarily theoretical in nature rather than giving a practical tracking algorithm. However, it does provide a method for starting a tracking algorithm from a minimal amount of data. We do not believe that people try to infer the motion of flying objects from as few snapshots as possible. They use all channels of data gathering, doing the job based on continuous and long sequences of stereo images provided by their eyes. They also most probably apply a tracking algorithm, continually guessing and predicting the motion of the object in free flight, and correcting their assumptions based on the differences between their predictions and actual observations. Hence to really build a robot with the capability of catching a flying object we should concentrate on continuous image input based estimation processes, like the extended Kalman filter. Such an algorithm should be capable not only of tracking the motions of the object but also of accounting for disturbances and effects due to air flow and friction encountered by the object in flight, without the need for complete knowledge of the physical laws governing the motion. Such robust tracking procedures are currently under consideration.

APPENDIX

In this section the proofs of the algebraic geometry theorems needed for Theorem 1 are given. Some terms are needed and will now be defined. See also [4, 5, 11].

On the complex n-space ($\mathbf{C}^n$), an (algebraic) variety is the zero set of finitely many polynomials. If V is a variety, then the Zariski topology on V is obtained by defining the closed sets of the topology to be the subvarieties (subsets of V which are themselves varieties) of V. Sets that are both open and dense in the Zariski topology are called Zariski open dense, and these are also both open and dense in the usual complex topology. A variety is irreducible if it cannot be expressed as the union of two proper closed subsets. The projective n-space $\mathbf{P}^n$ is the set of $(n+1)$-tuples $(z_0, z_1, \ldots, z_n)$ besides $(0, \ldots, 0)$ where points lying on the same line through the origin are identified. The notation $A\backslash B$ indicates the set of points in A which are not in B, and $\overline{A}$ denotes the closure of the set A. The proofs of the two lemmas below can be found in [4] and [5].

LEMMA 1. *Let V be an irreducible variety. Let f be a rational map from V to $\mathbf{C}^N$ given by $(p_1/q_1, \ldots, p_N/q_N)$, where the p_i and q_i are polynomials and in particular $\prod_{i=1}^N q_i$ is not identically 0 on V. Define $V' = \{\mathbf{z} \in V | \prod_{i=1}^N q_i(\mathbf{z}) = 0\}$. Then $f(V\backslash V')$ contains an irreducible variety which is Zariski open dense in $\overline{f(V\backslash V')}$.*

LEMMA 2. *Let $\mathbf{f}(\mathbf{z}, \mathbf{q})$ be a system of polynomials $f_i(\mathbf{z}, \mathbf{q})$, $i = 1, \ldots, N$, where $\mathbf{z} \in U = \prod_j \mathbf{P}^{n_j}$, a product of projective spaces, and $\mathbf{q} \in Q$ where Q is an irreducible variety. Suppose there is a $\mathbf{q}_0 \in Q$ such that $\mathbf{f}(\mathbf{z}, \mathbf{q}_0) = \mathbf{0}$ has only a finite number K of solutions. Then there is a Zariski open dense set $Q_0 \subseteq Q$ such that $\mathbf{f}(\mathbf{z}, \mathbf{q}) = \mathbf{0}$ has at most K solutions for all $\mathbf{q} \in Q_0$. If the dimension of U is N (that is, $\sum_j n_j = N$), then $\mathbf{f}(\mathbf{z}, \mathbf{q}) = \mathbf{0}$ has exactly K solutions for all $\mathbf{q} \in Q_0$.*

THEOREM 1. *Suppose we are given the perspective projections of a rod, in torque-free rigid motion about its unknown center of mass, at five evenly spaced time intervals. Suppose further that the path of the center of mass is known to be a conic section. Then there is in general a unique solution for the position and rotation of the rod.*

Proof. We will show that the system comprising Eqs. (11), (14)–(16) has exactly one real solution for $\mathbf{R}$, k, and all the z-coordinates of the feature points up to a common scale factor for almost all possible values of the observed data. This will be accomplished by showing that this system has two solutions in general. The only difference in the two solutions is that the signs of the quaternion variables $\{q_w, q_x, q_y, q_z\}$ are all reversed. With this information, we will know five points on the conic, which is sufficient to determine it uniquely, though again it is determined only up to the same common scale factor as the z-coordinates of the feature points.

In order to use Lemmas 1 and 2, a parameter space W must be introduced. Here the parameter space W will be the set of all points $(X_1, Y_1, X_2, Y_2, X_1', \ldots, Y_2^{(4)})$ in $\mathbf{C}^{20}$ such that there exists a 3×3 rotation matrix $\mathbf{R}$, 3×1 vectors $\mathbf{C}_0, \ldots, \mathbf{C}_4$, and a scalar k such that Eq. (2) holds, $j = 0, \ldots, 4$.

It will now be shown that the system of Eqs. (1) and (2) satisfies the hypotheses of Lemma 1. The irreducible va-

riety V is $(\mathbf{C}^3)^2 \times SO(3) \times (\mathbf{C}^3)^5 \times \mathbf{C}$, where $SO(3)$ is the special orthogonal group of 3×3 matrices with determinant 1. A point $\mathbf{w}$ in V can be regarded as a 13-tuple $(x_1, y_1, z_1, x_2, y_2, z_2, \mathbf{R}, \mathbf{C}_0, \mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3, \mathbf{C}_4, k)$. The closed set V' is $\{\mathbf{w} \in V | \prod_j z_1^{(j)} z_2^{(j)} = 0\}$, which eliminates degenerate cases in which a feature point lies on the camera plane. The map $f: V \backslash V' \rightarrow W$ is given by $f(x_1, y_1, z_1, x_2, y_2, z_2, \mathbf{R}, \mathbf{C}_0, \mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3, \mathbf{C}_4, k) = (X_1, Y_1, X_2, Y_2, X_1', \ldots, Y_2^{(4)})$, where the $X_i^{(j)}$, $Y_i^{(j)}$, and $z_i^{(j)}$ are given by Eqs. (2) and (1). The parameter space W is $f(V \backslash V')$. All the hypotheses of Lemma 1 are satisfied, so there is an irreducible variety $Q \subseteq W$ which is Zariski open dense in $\overline{W}$.

Now it will be shown that the system comprised of Eqs. (11), (14)–(16) satisfies the hypotheses of Lemma 2. All of these equations are homogeneous in the three sets of variables $\{q_0, q_w, q_x, q_y, q_z\}$, $\{z_1, z_2, z_1', \ldots, z_2^{(4)}\}$, $\{k, k_0\}$, so the product of projective spaces U of Lemma 2 is $\mathbf{P}^4 \times \mathbf{P}^9 \times \mathbf{P}^1$. The irreducible variety Q is the same as the one above guaranteed by Lemma 1. The $f_i(\mathbf{q}, \mathbf{z})$ is given by Eqs. (11, 14–16); there is a total of 24 of these. Since the number of equations exceeds the dimension of U, we can use the first part of the conclusion of Lemma 2. This tells us that if there is a finite number K of solutions to the system at one point of Q, then there are at most K solutions at a Zariski open dense subset of Q. Here K is equal to two. However, since Eqs. (11), (14)–(16) contain only terms of even degree in the variables $\{q_w, q_x, q_y, q_z\}$, (q_0 is excluded here), $(q_0, -q_w, -q_x, -q_y, -q_z)$ is a solution whenever $(q_0, q_w, q_x, q_y, q_z)$ is, and these correspond to the same rotation. An example with two solutions is that where $(X_1, Y_1, X_2, Y_2, X_1', \ldots, Y_2^{(4)})$ is given by Example 1. The two solutions in $\mathbf{P}^4 \times \mathbf{P}^9 \times \mathbf{P}^1$ may be expressed as $(q_0, q_w, q_x, q_y, q_z) = (3(10)^{1/2}, \pm 9, \pm 1, \mp 2, \pm 2)$, $(z_1, z_2, z_1', \ldots, z_2^{(4)}) = (75, 185, 150, 250, 300, 350, 240, 220, 219, 137)$, and $(k, k_0) = (1/4, 1)$. This result was obtained through the use of Macaulay, Maple, Macsyma, and Polsys. Several other examples tested also had exactly one solution for the position and rotation of the rod. By construction, every point of Q corresponds to a system with some solution. Therefore the system comprising Eqs. (11), (14)–(16) has exactly one solution at all points of a Zariski open dense subset of Q. ■

ACKNOWLEDGMENT

The authors express their gratitude to the referees for several useful comments.

REFERENCES

1. R. Bogen *et al.*, *MACSYMA Reference Manual*, Symbolics Incorporated, Cambridge, MA, 1983.
2. J. Canny, *The Complexity of Robot Motion Planning*, The MIT Press, Cambridge, MA, 1988.
3. B. W. Char, K. O. Geddes, G. H. Gonnet, M. B. Monagan, and S. M. Watt. *Maple V User's Guide*, Watcom, Waterloo, 1990.
4. R. J. Holt and A. N. Netravali, Motion and structure from multiple frame correspondence, AT&T Tech. Memo. 11256-900706.01TM, 1990.
5. R. J. Holt and A. N. Netravali, Motion from optic flow: Multiplicity of solutions, *J. Visual Comm. and Image Representation*, **4**, 1993, 14–24.
6. B. K. P. Horn, Closed-form solution of absolute orientation using unit quaternions, *J. Opt. Soc. Amer. A* **4**, 1987, 629–642.
7. T. S. Huang and A. N. Netravali, Motion and structure from feature correspondences: A review, *Proc. IEEE* **82**, 1994, 252–268.
8. Y. Liu and T. S. Huang, Estimation of rigid body motion using straight line correspondences, *Comput. Vision, Graphics, and Image Process.* **43**, 1988, 37–52.
9. Y. Liu and T. S. Huang, A linear algorithm for motion estimation using straight line correspondences, *Comput. Vision, Graphics, and Image Process.* **44**, 1988, 35–57.
10. S. W. McCuskey, *Introduction to Advanced Dynamics*, Addison-Wesley, Reading, MA, 1959.
11. A. P. Morgan and A. J. Sommese, Coefficient-parameter polynomial continuation, *Appl. Math. Comput.* **29**, 1989, 123–160.
12. A. P. Morgan, A. J. Sommese, and L. T. Watson, Finding all solutions to polynomial systems using HOMPACK, *ACM Trans. on Math. Software* **15**, 1989, 93–122.
13. H. Shariat and K. Price, The motion problem: How to use more than two frames. *IEEE Trans. Pattern Anal. Mach. Intell.* **12**, 1990, 417–434.
14. M. Stillman, M. Stillman, and D. Bayer, *Macaulay User's Manual*, Version 3.0, 1989.
15. B. L. Van der Waerden, *Modern Algebra*, Unger, New York, 1950.

ALFRED M. BRUCKSTEIN received the B.Sc. and M.Sc. degrees in electrical engineering from the Technion, Israel Institute of Technology, in 1977 and 1980, respectively, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, California, in 1984. Since October 1984 he has been with the Technion, Haifa, Israel. He is a frequent visitor at AT&T Bell Laboratories, Murray Hill, New Jersey. His present research interests are computer vision, pattern recognition, image processing, and computer graphics. He has also done work in estimation theory, signal processing, algorithmic aspects of inverse scattering, point processes and mathematical models in neurophysiology. Professor Bruckstein is a member of SIAM, MAA, and AMS.

ROBERT J. HOLT received the B.S. degree in mathematics from Standford University, Stanford, California, in 1982 and the Ph.D. De-

gree in mathematics from the Massachusetts Institute of Technology, Cambridge, Massachusetts, in 1986. He was Assistant Professor of Mathematics at Northwestern University, Evanston, Illinois, from 1986 to 1988. Since then he has been working at AT&T Bell Laboratories in Murray Hill, New Jersey, and is currently in the Interactive Computing Systems Research Department. His research interests include computer vision and image processing.

ARUN NETRAVALI received the B. Tech. (Honors) degree from the Indian Institute of Technology, Bombay, India, in 1967 and the M.S. and Ph.D. degrees from Rice University, Houston, Texas, in 1969 and 1970, respectively, all in electrical engineering. Dr. Netravali is currently the Executive Director of Research of the Communications Sciences Division and Technology Conversion Laboratories at AT&T Bell Laboratories, with responsibility for research in all aspects of communication and networking. He has also been an adjunct professor at the Massachusetts Institute of Technology since 1984, and is currently an editor of several journals. Dr. Netravali is a member of the Tau Beta Pi and Sigma Xi, a Fellow of the IEEE and AAAS, and a member of the United States National Academy of Engineering. He is the author of more than 100 papers and holds over 50 patents in the areas of computer networks, human interfaces to machines, picture processing, and digital television. He is the co-author of two books: (a) "Digital Picture Representation and Compression" (Plenum, 1987), and (b) "Visual Communication Systems" (IEEE Press, 1989). He has recently received the Alexander Graham Bell Medal in 1991, the OCA National Corporate Employee Achievement Award in 1991, and the Engineer of the Year Award from the Association of Engineers from India in 1992. He serves on the New Jersey Governor's Committee on "Schools" programs.