Biol. Cybern. 54, 71–83 (1986)

Biological Cybernetics

The Computation of Structure from Fixed-Axis Motion: Rigid Structures

D. D. Hoffman[1] and B. M. Bennett[2]

[1] School of Social Sciences, University of California, Irvine, CA 92717, USA
[2] Department of Mathematics, University of California, Irvine, CA 92717, USA

Abstract. We show that three distinct orthographic views of three points in a rigid configuration are compatible with at most 64 interpretations of the three-dimensional structure and motion of the points. If, in addition, one assumes that the three points spin about a fixed axis over the three views, then with probability one there is a unique three-dimensional interpretation (plus a reflection). Moreover the probability of false targets is zero. In the special case that the axis of rotation is parallel to the image plane three views of the three points are sufficient to obtain at most two interpretations (plus reflections) – unless one assumes the angular velocity about the axis is constant, in which case three views of two points are sufficient to determine a unique interpretation. Closed form solutions are obtained for each of these cases. The systems of equations studied here are in each case overconstraining (i.e. there are more independent equations than unknowns) and are amenable to solution by nonlinear programming. These two properties make possible the construction of noise insensitive algorithms for computer vision systems. Our uniqueness proofs employ the *principle of upper semicontinuity*, a principle which underlies a general mathematical framework for the analysis of solutions to overconstraining systems of equations.

1 Introduction

Valuable information is lost in the projection from the visible environment onto the human retina. For instance, all points in the environment along a line of sight project to a single point on the retina. In consequence the retinal image is but two-dimensional whereas the visible environment is three-dimensional. Yet most observers perceive the world as three-dimensional, unaware that the dimension of depth is unavailable to the eye and therefore must be reconstructed.

Motion is one means used by the visual system for the reconstruction (Braunstein 1976; Gibson 1950; Helmholtz 1925). Psychophysical experiments show that one can perceive the three-dimensional structure of an object from its changing retinal projections even when the structure is unfamiliar and a static view of the object gives no perception of depth (Wallach and O'Connell 1953).

Retinal motion is, however, insufficient by itself to determine uniquely the three-dimensional structure of the environment. In fact an infinite number of three-dimensional interpretations are always equally compatible with the changing retinal image because motion along the line of sight is lost in the retinal projection. As a result, to make possible a unique three-dimensional interpretation from retinal motion, certain regularities in the visual world must be exploited by the visual system. These regularities are roughly of two types: structural and dispositional. *Structural* regularities are regularities in the motion of the points of an object relative to each other. *Dispositional* regularities are regularities in the motions of the points of an object relative to an external frame of reference (e.g., the frame of reference of an observer).

One plausible structural regularity is rigidity: all points of a rigid object move relative to each other so as to maintain a constant distance (Gibson and Gibson 1957; Green 1961; Hay 1966; Hoffman 1983; Johansson 1975; Reuman and Hoffman 1986; Ullman 1979; Wallach and O'Connell 1953). Ullman (1979) proved that using this regularity alone one can in principle recover three-dimensional structure and motion from three orthographic (parallel) projections of four or more points. Longuet-Higgins and Prazdny (1980) proved that the first and second spatial derivatives of the perspectively projected velocity field of a rigid object determine the surface normal at each point and the relative motion. Hoffman (1982) showed that if the motion field is viewed under orthographic projection (rather than perspective) then the spatial derivatives of the acceleration field are also required in order to recover the structure of rigid objects.

Because many visual objects are not rigid (Johansson 1975), structural regularities other than rigidity must be explored. Bennett and Hoffman (1985) investigated a common axis regularity: the points of a common axis object move in parallel circles having collinear centers, but do so at independent angular velocities. They found that four orthographic views of two points or three orthographic views of four points are compatible with at most one three dimensional interpretation (plus its reflection).

Dispositional regularities are lawful properties of the motion of an object relative to an external observer. For instance, one frequently observes fixed-axis motion (Webb and Aggarwal 1982): the points of an object spin about an axis whose orientation does not change with respect to the observer. One common case, discussed in Sect. 4, occurs when an observer is translating in a straight line through a static environment. If the observer foveates a point in the environment as he translates then all points in the environment undergo an induced fixed-axis motion about an axis that is orthogonal to the observer's line of sight and direction of motion. Bobick (1983) has shown that if one uses the fixed-axis regularity then two views of three points together with velocity direction vectors at each point are compatible with at most one three-dimensional interpretation (plus an orthographic reflection). Hoffman and Flinchbaugh (1982) have shown that if one uses the more restrictive assumption of planar motion then three views of two points or two views of three points are compatible with at most one three-dimensional interpretation (plus reflection).

It appears that regularities of structure and disposition are both used by the human visual system. If, for example, rigidity alone were used by the visual system then one would expect that observers could perceive the structure of rigid objects whose motion was quite jerky. In fact such displays give poor impressions of depth. Again, if rigidity alone were used one would expect that observers could not perceive three-dimensional structure in displays having fewer than three views or four rigid points (as Ullman's result requires). However observers can see the three-dimensional structure in many such displays, for instance in the biological motion displays of Johansson (1975).

In this paper we determine conditions under which one can recover the three-dimensional structure of objects using the structural regularity of rigidity in combination with the dispositional regularity of fixed-axis motion. In Sect. 2 we show that three orthographic views of three points in a rigid configuration are compatible with at most two interpretations of the three-dimensional structure (plus reflections) for the first view. However there are sixty four possible motions for the structures over the three views. We give simple closed form solutions for the interpretations. In Sect. 3 we show that by adding a fixed-axis constraint one eliminates all but one of the global solutions obtained in Sect. 2 (plus a reflection). In addition one eliminates "false targets", points in three dimensions not undergoing rigid fixed-axis motion but whose projections appear consistent with such motion. The proofs of uniqueness and no false targets use semicontinuity theorems from algebraic geometry, theorems that allow proof of uniqueness and no false targets by a single concrete example in each case.

A degenerate case for the analysis of Sect. 3 occurs when the axis of rotation is parallel to the image plane. In Sect. 4 we show that this is an important special case. A translating observer who foveates a point induces a rotary motion of the environment about a fixed axis that is always parallel to the image plane. In Sect. 5 we provide closed form solutions for the special case when the axis is parallel to the image plane and the points move at constant angular velocity about the axis. In this case three views of two points are sufficient for a unique interpretation, and the probability of false targets is zero. In Sect. 6 we provide closed form solutions for the case when the angular velocity is not necessarily constant. Three views of three points are compatible with at most two interpretations in this analysis, and again the probability of false targets is zero.

Pilot studies by Braunstein (personal communication) indicate that human observers can in fact recover the three-dimensional structure of rigid bodies in fixed-axis motion from as few as three views of three points.

2 Rigid Structures

In this section we prove the following result:

Theorem 2.0. *Given three generic orthographic views of three points in a rigid configuration, there are two interpretations of the three-dimensional structure (plus orthographic reflections) for the first view. There are sixteen possible motions for each structure, giving a total of sixty four motions.*

The system of equations studied in this section is not overconstraining. In fact, the number of independent equations is equal to the number of unknowns (six). The next section (Sect. 3) adds two overconstraining equations which arise from the constraint of fixed-axis motion. Section 3 also discusses a general mathematical framework for analyzing the solutions to overconstrained systems of equations.

Proof. Call the three points O, A_1, and A_2. Let $\mathbf{a}_{ij}$ be the vector (in three dimensions) between O and point A_i in view j ($j=1,2,3$) as shown in Fig. 1. Because the three points are in a rigid configuration we expect that the length of the vector from O to A_1 remains constant over all three views. Similarly we expect that the length of the vector from O to A_2 remains constant over all three views. Consequently we can write

$$\mathbf{a}_{11}\cdot\mathbf{a}_{11}=\mathbf{a}_{12}\cdot\mathbf{a}_{12}, \tag{2.1a}$$

$$\mathbf{a}_{11}\cdot\mathbf{a}_{11}=\mathbf{a}_{13}\cdot\mathbf{a}_{13}, \tag{2.1b}$$

$$\mathbf{a}_{21}\cdot\mathbf{a}_{21}=\mathbf{a}_{22}\cdot\mathbf{a}_{22}, \tag{2.1c}$$

$$\mathbf{a}_{21}\cdot\mathbf{a}_{21}=\mathbf{a}_{23}\cdot\mathbf{a}_{23}. \tag{2.1d}$$

In addition we expect that the angle between the vector OA_1 and vector OA_2 remains constant over all three views. Thus we can write

$$\mathbf{a}_{11}\cdot\mathbf{a}_{21}=\mathbf{a}_{12}\cdot\mathbf{a}_{22}, \tag{2.2a}$$

$$\mathbf{a}_{11}\cdot\mathbf{a}_{21}=\mathbf{a}_{13}\cdot\mathbf{a}_{23}. \tag{2.2b}$$

To solve these six equations it is useful to express the $\mathbf{a}_{ij}$'s in terms of components. Let $\mathbf{a}_{ij}=(x_{ij}, y_{ij}, z_{ij})$. Assume that the line of sight lies along the z-axis. Then the x_{ij}'s and y_{ij}'s are known directly from the views. The six z_{ij}'s are unknown and must be solved for.

Equation (2.1) may be expressed in terms of components as

$$z_{11}^2-z_{12}^2+c_1=0, \tag{2.3a}$$

$$z_{11}^2-z_{13}^2+c_2=0, \tag{2.3b}$$

$$z_{21}^2-z_{22}^2+c_3=0, \tag{2.3c}$$

$$z_{21}^2-z_{23}^2+c_4=0. \tag{2.3d}$$

Equation (2.2) may be expressed in terms of components as

$$z_{11}z_{21}-z_{12}z_{22}+c_5=0, \tag{2.4a}$$

$$z_{11}z_{21}-z_{13}z_{23}+c_6=0, \tag{2.4b}$$

where

$$c_1=x_{11}^2+y_{11}^2-x_{12}^2-y_{12}^2, \tag{2.5a}$$

$$c_2=x_{11}^2+y_{11}^2-x_{13}^2-y_{13}^2, \tag{2.5b}$$

$$c_3=x_{21}^2+y_{21}^2-x_{22}^2-y_{22}^2, \tag{2.5c}$$

$$c_4=x_{21}^2+y_{21}^2-x_{23}^2-y_{23}^2, \tag{2.5d}$$

$$c_5=x_{11}x_{21}+y_{11}y_{21}-x_{12}x_{22}-y_{12}y_{22}, \tag{2.5e}$$

$$c_6=x_{11}x_{21}+y_{11}y_{21}-x_{13}x_{23}-y_{13}y_{23}. \tag{2.5f}$$

For the convenience of the reader we describe one way to solve these equations: Use (2.3a) and (2.3c) to

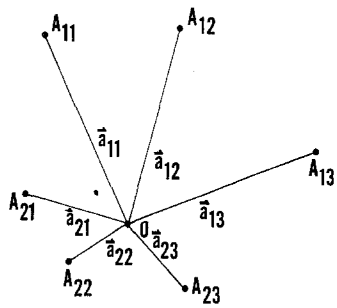

Fig. 1. Geometry underlying the computation of structure from three orthographic views of three points in a rigid configuration

eliminate z_{12} and z_{22}, respectively, from (2.4a). Use (2.3b) and (2.3d) to eliminate z_{13} and z_{23}, respectively, from (2.4b):

$$z_{11}z_{21}\pm\sqrt{(z_{11}^2+c_1)(z_{21}^2+c_3)}+c_5=0, \tag{2.6a}$$

$$z_{11}z_{21}\pm\sqrt{(z_{11}^2+c_2)(z_{21}^2+c_4)}+c_6=0. \tag{2.6b}$$

The $\pm$ before the radicals in (2.6) indicates that these equations may be rewritten as

$$(z_{11}z_{21}+\sqrt{(z_{11}^2+c_1)(z_{21}^2+c_3)}+c_5)(z_{11}z_{21}-\sqrt{(z_{11}^2+c_1)(z_{21}^2+c_3)}+c_5)=0, \tag{2.7a}$$

$$(z_{11}z_{21}+\sqrt{(z_{11}^2+c_2)(z_{21}^2+c_4)}+c_6)(z_{11}z_{21}-\sqrt{(z_{11}^2+c_2)(z_{21}^2+c_4)}+c_6)=0. \tag{2.7b}$$

Expand and simplify (2.7):

$$c_3z_{11}^2+c_1z_{21}^2-2c_5z_{11}z_{21}+c_1c_3-c_5^2=0, \tag{2.8a}$$

$$c_4z_{11}^2+c_2z_{21}^2-2c_6z_{11}z_{21}+c_2c_4-c_6^2=0. \tag{2.8b}$$

[Note that the fourth degree terms cancel to make (2.8a) and (2.8b) of only second degree.] Let $c_7=c_1c_3-c_5^2$ and $c_8=c_2c_4-c_6^2$. Multiply (2.8a) by c_8. Multiply (2.8b) by c_7. Subtract (2.8b) from (2.8a) and simplify to obtain

$$(c_3c_8-c_4c_7)z_{11}^2+2(c_6c_7-c_5c_8)z_{11}z_{21}+(c_1c_8-c_2c_7)z_{21}^2=0. \tag{2.9}$$

Let $\zeta=z_{11}/z_{21}$, $a=c_3c_8-c_4c_7$, $b=2(c_6c_7-c_5c_8)$, and $c=c_1c_8-c_2c_7$. Divide (2.9) by z_{21}^2 to get

$$a\zeta^2+b\zeta+c=0. \tag{2.10}$$

The quantity a is in general not zero as may be verified directly from the definition of the c_i's. Thus, since we assume generic views, we may solve (2.10) for ζ:

$$\zeta=\frac{-b\pm\sqrt{b^2-4ac}}{2a}. \tag{2.11}$$

Rewrite (2.8a) using ζ:

$$c_3\zeta^2z_{21}^2+c_1z_{21}^2-2c_5\zeta z_{21}^2+c_7=0. \tag{2.12}$$

Solve (2.12) for z_{21}:

$$z_{21} = \pm \sqrt{\frac{-c_7}{c_3\zeta^2 - 2c_5\zeta + c_1}}. \tag{2.13}$$

Now it is an easy matter to obtain all the remaining z_{ij}'s. Find z_{11} using the relation $z_{11} = z_{21}\zeta$. Find z_{12} using (2.3a), z_{13} using (2.3b), z_{22} using (2.3c), and z_{23} using (2.3d). The genericity hypothesis in the statement of (2.0) corresponds to the data $\{(x_{ij}, y_{ij})\}$ satisfying both $a \neq 0$ and $c_3\zeta^2 - 2c_5\zeta + c_1 \neq 0$.

Observe that we have shown there are but two interpretations (plus reflections) for the structure in the first frame, and sixty four possible motions for the structures. In the first frame, for example, (2.13) gives at most four solutions for z_{21}, two solutions of opposite sign for each of the two values of ζ. To each z_{21} there is associated a unique value of z_{11} by the equation $z_{11} = \zeta z_{21}$. Consequently there are a total of two interpretations (plus reflections) for the first frame. For the second frame, (2.3a) gives four solutions for z_{12} and (2.3c) gives four solutions for z_{22}, again two solutions of opposite sign for each of the two values of ζ. Note that the choice of sign for z_{11} does not determine the choice of sign for z_{12}, since (2.3a) is an equation involving only the squares of these two variables. Similarly, the choice of sign for z_{21} does not determine the choice of sign for z_{22}. The same is true for z_{13} and z_{23} in the third frame. The result is that there are two structures plus reflections, each having 4^2 (sixteen) possible motions over the three frames (the different motions arising because the choice of reflection in one frame does not determine the choice in succeeding frames). This gives a total of sixty four motions.

3 Fixed-Axis Motion: Generic Case

In Sect. 2 we conclude that three views of three points in a rigid configuration does not, in general, allow a unique three-dimensional interpretation, but does reduce the possible interpretations to two structures in a total of sixty four motions. We should also note that there are no more equations than unknowns so that the probability of false targets (nonrigid objects giving rise to projections compatible with a rigid interpretation) is greater than zero.

There are at least three ways to take the result of Sect. 2 one step further to eliminate false targets and extra interpretations. One could add a fourth point or add a fourth view or add a dispositional constraint. Adding a fourth point leads to Ullman's (1979) result that three views of four points give a unique interpretation. In this section we add instead the dispositional constraint of fixed-axis motion and prove the following:

Theorem 3.0 (i) *Given three orthographic projections of three points spinning rigidly about a fixed axis, the probability is one that the three-dimensional structure and motion of the points is uniquely determined (up to a reflection about the image plane). Moreover* (ii) *the probability is zero that a randomly chosen set of image data permits such a determination.*

To prove this we first introduce two linear equations that express the fixed-axis constraint. These equations, together with (2.3) and (2.4), will give us eight equations in six unknown z_{ij}'s, whose coefficients depend on six pairs of image coordinates (x_{ij}, y_{ij}). We will prove that this system of equations has no solutions for generic choices of (x_{ij}, y_{ij}), thus demonstrating that the probability of false targets is zero – the second assertion of Theorem 3.0 above. We will also show that among those (x_{ij}, y_{ij}) for which the equations admit at least one solution, the condition that the solution is unique (up to reflection) is generic – the first assertion of Theorem 3.0.

One means of expressing the fixed-axis motion constraint is illustrated in Fig. 2. As can be seen from the figure, fixed-axis motion implies that the difference vectors between the different positions of the first point must be coplanar with the difference vectors between the different positions of the second point. We can write that the scalar triple product of three of these difference vectors is zero:

$$(\mathbf{a}_{11} - \mathbf{a}_{12}) \cdot [(\mathbf{a}_{11} - \mathbf{a}_{13}) \times (\mathbf{a}_{21} - \mathbf{a}_{22})] = 0, \tag{3.1a}$$

$$(\mathbf{a}_{11} - \mathbf{a}_{12}) \cdot [(\mathbf{a}_{11} - \mathbf{a}_{13}) \times (\mathbf{a}_{21} - \mathbf{a}_{23})] = 0. \tag{3.1b}$$

Expanding (3.1) in terms of components gives

$$a_1 z_{11} + a_2 z_{12} + a_3 z_{13} + a_4 z_{21} + a_5 z_{22} = 0, \tag{3.2a}$$

$$a_6 z_{11} + a_7 z_{12} + a_8 z_{13} + a_4 z_{21} + a_5 z_{23} = 0, \tag{3.2b}$$

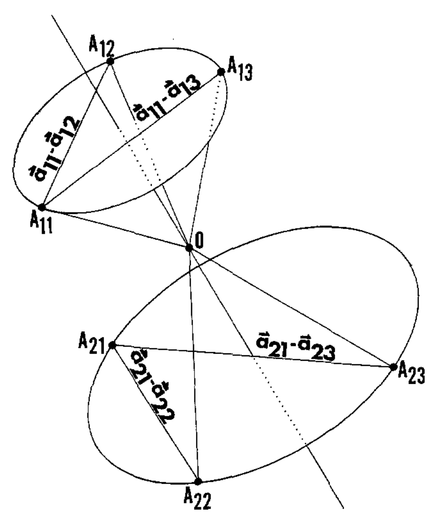

Fig. 2. Fixed axis motion implies coplanarity of the indicated difference vectors

where

$$a_1=(x_{12}-x_{13})(y_{21}-y_{22})-(x_{21}-x_{22})(y_{12}-y_{13}),$$
$$a_2=(x_{21}-x_{22})(y_{11}-y_{13})-(x_{11}-x_{13})(y_{21}-y_{22}),$$
$$a_3=(x_{11}-x_{12})(y_{21}-y_{22})-(x_{21}-x_{22})(y_{11}-y_{12}),$$
$$a_4=(x_{11}-x_{12})(y_{11}-y_{13})-(x_{11}-x_{13})(y_{11}-y_{12}),$$
$$a_5=-a_4, \qquad (3.3)$$
$$a_6=(x_{12}-x_{13})(y_{21}-y_{23})-(x_{21}-x_{23})(y_{12}-y_{13}),$$
$$a_7=(x_{21}-x_{23})(y_{11}-y_{13})-(x_{11}-x_{13})(y_{21}-y_{23}),$$
$$a_8=(x_{11}-x_{12})(y_{21}-y_{23})-(x_{21}-x_{23})(y_{11}-y_{12}).$$

We summarize: *Let* $\{(x_{ij}, y_{ij}, z_{ij})\}_{\substack{i=1,2\\ j=1,2,3}}$ *be the coordinates in* $\mathfrak{R}^3$ *of three successive positions* $(j=1,2,3)$ *of each of two points* $(i=1,2)$. *These positions are compatible with an interpretation that the points are spinning rigidly about a fixed axis through the origin if and only if the coordinates satisfy the eight equations*

$$z_{11}^2-z_{12}^2+c_1=0, \qquad (3.4a)$$
$$z_{11}^2-z_{13}^2+c_2=0, \qquad (3.4b)$$
$$z_{21}^2-z_{22}^2+c_3=0, \qquad (3.4c)$$
$$z_{21}^2-z_{23}^2+c_4=0, \qquad (3.4d)$$
$$z_{11}z_{21}-z_{12}z_{22}+c_5=0, \qquad (3.4e)$$
$$z_{11}z_{21}-z_{13}z_{23}+c_6=0, \qquad (3.4f)$$
$$a_1z_{11}+a_2z_{12}+a_3z_{13}+a_4z_{21}+a_5z_{22}=0, \qquad (3.4g)$$
$$a_6z_{11}+a_7z_{12}+a_8z_{13}+a_4z_{21}+a_5z_{23}=0, \qquad (3.4h)$$

where the coefficients $c_1, \ldots, c_6$ *and* $a_1, \ldots, a_8$ *are the polynomials in* x_{ij}, y_{ij} *defined in* (2.5) *and* (3.3). *The* (x_{ij}, y_{ij})*'s, and hence these coefficients, are accessible to an orthographic viewer whose line of sight is along the z-axis. Thus: three orthographic views* $\{(x_{ij}, y_{ij})\}$ *of two points are compatible with an interpretation of rigid motion about a fixed axis through the origin in* $\mathfrak{R}^3$ *if and only if the equations* (3.4) *(in which* $c_1, \ldots, c_6$ *and* $a_1, \ldots, a_8$ *are obtained from the particular viewing data* $\{(x_{ij}, y_{ij})\}$*) have a solution in the* z_{ij}*; each such solution corresponds to one possible interpretation.*

The techniques that we use to prove Theorem 3.0 require that we work temporarily with complex numbers, so that we will assume for the moment that the $\{(x_{ij}, y_{ij})\}$ can be complex. Thus $\{(x_{ij}, y_{ij}, z_{ij})\}_{i=1,2;j=1,2,3}$ is a point of $\mathbb{C}^{18}$ and $\{(x_{ij}, y_{ij})\}_{i=1,2;j=1,2,3}$ is a point of $\mathbb{C}^{12}$. Let q be the projection from $\mathbb{C}^{18}$ to $\mathbb{C}^{12}$, defined by $q(\{(x_{ij}, y_{ij}, z_{ij})\}) = \{(x_{ij}, y_{ij})\}$. Note that if $P=\{(x_{ij}, y_{ij})\}$ is a point of $\mathbb{C}^{12}$, then $q^{-1}(P)$ (the inverse image of P by q) is a copy of $\mathbb{C}^6$ with coordinates z_{ij}. We can interpret the solutions in the z_{ij} of the set of equations (3.4) (whose coefficients $a_1, \ldots, a_8$ and $c_1, \ldots, c_6$ are determined by the given P) as lying in that particular copy of $\mathbb{C}^6$. *For each* P, *let* $N(P)$ *denote the number of such solutions, counted with multiplicity;* $N(P)$ *may be infinity.* For each integer $m \geqq 0$, define

$$T_m=\{P \in \mathbb{C}^{12} \mid N(P) \geqq m\}.$$

Clearly

$$\mathbb{C}^{12}=T_0 \supset T_2 \supset \ldots \supset T_m \supset T_{m+1} \supset \ldots.$$

We can express things geometrically as follows: Remembering that $c_1, \ldots, c_6, a_1, \ldots, a_8$ are polynomials in the x_{ij} and y_{ij}, we view (3.4) as a system of 8 equations in 18 complex variables. The solutions of this system are therefore a locus W in $\mathbb{C}^{18}$. Then for $P \in \mathbb{C}^{12}$, $N(P)$ is the number of points (counted with multiplicities) in $q^{-1}(P) \cap W$, i.e. the number of points of W which lie over P, with respect to the projection $q: \mathbb{C}^{18} \to \mathbb{C}^{12}$. In particular $q(W)$ (the image of W in $\mathbb{C}^{12}$ by q) is the same as T_1. We will also use the letter V to denote this set.

$$\begin{array}{ccc} \mathbb{C}^{18} \supset & W \\ \downarrow q & \downarrow \\ \mathbb{C}^{12} \supset & V = T_1 \end{array}$$

An inspection of the equations (3.4) reveals that for any values of $c_1, \ldots, c_6, a_1, \ldots, a_8$ – i.e. for any $P \in \mathbb{C}^{12}$ – if (z_{ij}) is a solution then so is $(-z_{ij})$; this sign change corresponds to a "reflection about the image plane". Thus $V=T_1=T_2$, $T_3=T_4$, etc.

We are ultimately interested only in points with real number coordinates. Let $\mathfrak{R}^{12}$ and $\mathfrak{R}^{18}$ denote the subsets of $\mathbb{C}^{12}$ and $\mathbb{C}^{18}$ consisting of those points with real coordinates. Then we define

$$W(\mathfrak{R})=W \cap \mathfrak{R}^{12}$$
$$T_m(\mathfrak{R})=\{P \in T_m \cap \mathfrak{R}^{12} \mid q^{-1}(P) \cap W(\mathfrak{R}) \text{ contains at least } m \text{ points}\}.$$

In particular, we have

$$\mathfrak{R}^{12}=T_0(\mathfrak{R}) \supset T_2(\mathfrak{R})=V(\mathfrak{R}) \supset T_4(\mathfrak{R}),$$

etc. Note that in general a point P in $T_m \cap \mathfrak{R}^{12}$ may not be in $T_m(\mathfrak{R})$, since a priori some of the points in $W \cap q^{-1}(P)$ may not be in $W(\mathfrak{R})$, i.e. some solutions (z_{ij}) involving complex coordinates may contribute to the number $N(P)$ even though P itself is real.

We may now state succinctly the assertions of Theorem 3.0.

(3.5) (i) $T_4(\mathfrak{R})$ *has measure zero in* $T_2(\mathfrak{R})=V(\mathfrak{R})$ *with respect to any "unbiased" measure on* $V(\mathfrak{R})$. (ii) $T_2(\mathfrak{R})$ *has measure zero in* $\mathfrak{R}^{12}$.

Our technique of proof for both (i) and (ii) is based on the *principle of upper semicontinuity* which may be stated for our purposes as follows: Let S be a system of algebraic equations in projective space of arbitrary dimension over the complex numbers. Suppose that the coefficients of the equations in S depend algebraically (i.e. are polynomials) in some parameters which vary in $\mathbb{C}^n$. Then the function which assigns to each point P in $\mathbb{C}^n$ (i.e. to each choice of parameter values) the number of solutions counted with multiplicities to the system S, is upper semicontinuous in the Zariski topology on $\mathbb{C}^n$.

Now in the Zariski topology, by definition, the closed sets are closed algebraic varieties, i.e. solution sets to polynomial equations. Recall moreover that a function is upper semicontinuous if the locus of points where its value is greater than or equal to some given number is a closed set. Thus the upper semicontinuity principle translates into the following: Given any integer m, the set of parameter values in $\mathbb{C}^n$ for which the system admits at least m solutions is itself the solution set of a family of polynomial equations in $\mathbb{C}^n$. The interest for us here is that such a set is of strictly smaller dimension than $\mathbb{C}^n$ so that it has Lebesgue measure 0 in $\mathbb{C}^n$, or it is equal to $\mathbb{C}^n$ (in the case when the polynomials defining it are identically 0).

We will first prove assertion (ii) of (3.5). We would like to apply the upper semicontinuity principle directly to our system (3.4), to deduce that the function $N(P)$ defined above is upper semicontinuous on $\mathbb{C}^{12}$. Unfortunately we cannot do this directly because the principle only applies to systems of equations in complex projective space, whereas (3.4) is a system in complex affine space $\mathbb{C}^6$ (for any parameter choice P in $\mathbb{C}^{12}$). We can remedy this defect by canonically extending the system (3.4) into a system in complex projective space $\mathbb{P}^6$ as follows:

Set $z_{ij}=Z_{ij}/W$ where Z_{ij} and W are homogeneous coordinates on $\mathbb{P}^6$ (seven coordinates in all). We then obtain the extended system by multiplying by W^2.

$$Z_{11}^2-Z_{12}^2+c_1W^2=0, \qquad (3.6a)$$
$$Z_{11}^2-Z_{13}^2+c_2W^2=0, \qquad (3.6b)$$
$$Z_{21}^2-Z_{22}^2+c_3W^2=0, \qquad (3.6c)$$
$$Z_{21}^2-Z_{23}^2+c_4W^2=0, \qquad (3.6d)$$
$$Z_{11}Z_{21}-Z_{12}Z_{22}+c_5W^2=0, \qquad (3.6e)$$
$$Z_{11}Z_{21}-Z_{13}Z_{23}+c_6W^2=0, \qquad (3.6f)$$
$$a_1Z_{11}+a_2Z_{12}+a_3Z_{13}+a_4Z_{21}+a_5Z_{22}=0, \qquad (3.6g)$$
$$a_6Z_{11}+a_7Z_{12}+a_8Z_{13}+a_4Z_{21}+a_5Z_{23}=0, \qquad (3.6h)$$

The locus $W=0$ in $\mathbb{P}^6$ corresponds to the points added "at infinity" to $\mathbb{C}^6$ in order to form $\mathbb{P}^6$. Thus the solutions to (3.6) which do not correspond to solutions of (3.4) are those nontrivial solutions for which $W=0$. Now if $W=0$, the first four equations of (3.6) yield $Z_{12}=\pm Z_{11}$, $Z_{13}=\pm Z_{11}$, $Z_{12}=\pm Z_{21}$, $Z_{23}=\pm Z_{21}$. When we substitute these values in the last two equations we will get several systems of two homogeneous equations in Z_{11} and Z_{21} which will have a nontrivial solution if and only if the two equations are dependent. The condition for this dependence is that certain 2×2 determinants built out of $a_1, \ldots, a_8$ vanish. Since these may be expressed as polynomials in $\{(x_{ij}, y_{ij})\}$ we find: *There is a Zariski closed set C in $\mathbb{C}^{12}$ such that for $P\in\mathbb{C}^{12}-C$ the extended system* (3.6) *has no more solutions than* (3.4), *i.e.* (3.6) *has no solutions "at infinity"*.

Now let us apply the upper semicontinuity principle to the system (3.6). It tells us in particular that the set B in $\mathbb{C}^{12}$ where the number of solutions to (3.6) is at least 1 is Zariski closed. Since every solution to (3.4) is also a solution to (3.6) it is clear that $V\subset B$. Thus if we can show B is a *proper* Zariski closed subset of $\mathbb{C}^{12}$ (i.e. $B\neq\mathbb{C}^{12}$) it follows that it has Lebesgue measure zero in $\mathbb{C}^{12}$. In other words, because of the upper semicontinuity principle, to show B has measure zero it suffices to find *one point* of $\mathbb{C}^{12}$ not in it. As the authors have done, the reader may select a value at random for the point $P=\{(x_{ij}, y_{ij})\}\in\mathbb{C}^{12}$, and verify that the resulting equations (3.6) have no solution. We remark that solutions to (3.6) are of two types, namely those which are solutions to (3.4), and those which are not, the latter being the nontrivial solutions to (3.6) with $W=0$, i.e. the solutions "at infinity". This corresponds to the fact that $B=V\cup C$.

Since $V\subset B$, it follows that V also has measure zero in $\mathbb{C}^{12}$. However this is not yet our desired conclusion; we want to show that $V(\mathfrak{R})$ $(=T_1(\mathfrak{R})=T_2(\mathfrak{R}))$ has measure zero in $\mathfrak{R}^{12}$. Since $V(\mathfrak{R})\subset V\cap\mathfrak{R}^{12}\subset B\cap\mathfrak{R}^{12}$, it suffices to show that $B\cap\mathfrak{R}^{12}$ has measure zero in $\mathfrak{R}^{12}$. Indeed we note that *any* proper closed algebraic subvariety B of $\mathbb{C}^n$ intersects $\mathfrak{R}^n$ in a measure zero subset. To see this, first observe that by hypothesis on B, there is a non-zero polynomial $f(p_1, \ldots, p_n)$ which vanishes on B where $p_i=r_i+\sqrt{-1}q_i$ are the complex coordinates on $\mathbb{C}^n$. Since $\mathfrak{R}^n$ is the subset of $\mathbb{C}^n$ where the q_i's vanish, $B\cap\mathfrak{R}^n$ is the solution set in $\mathfrak{R}^n$ of the polynomial $f(r_1, \ldots, r_n)$, which is a polynomial in the n real variables $r_1, \ldots, r_n$, but which has nonzero complex coefficients. By decomposing the coefficients into their real and complex parts, we can write $f(r_1, \ldots, r_n)=g(r_1, \ldots, r_n)+\sqrt{-1}h(r_1, \ldots, r_n)$ where now g and h are real polynomials, at least one of which has nonzero coefficients, i.e. is not identically zero on $\mathfrak{R}^n$. These polynomials must both vanish on $B\cap\mathfrak{R}^n$. Now suppose that $B\cap\mathfrak{R}^n$ did not have measure zero in $\mathfrak{R}^n$. Then there would be an open set U of $\mathfrak{R}^n$ in which $B\cap\mathfrak{R}^n$ is dense. Since g and h vanish on B, and since polynomials are continuous functions, they vanish on all of U.

Moreover, since polynomials are (real) analytic functions on $\mathfrak{R}^n$ they are completely determined by their value on any open set of $\mathfrak{R}^n$. Hence both g and h are identically zero on $\mathfrak{R}^n$, a contradiction. This completes the proof of (3.5), (ii).

We now proceed to the proof of (3.5 (i). We can produce a point of $V(\mathfrak{R})$ which is not in T_4 (see below). By the upper semicontinuity principle, it will follow that T_4 is a *proper* closed algebraic subvariety of $V=T_2$. We cannot conclude directly in this case, however, that T_4 has measure zero in $T_2=V$ (or what is more, that $T_4(\mathfrak{R})$ has measure zero in $V(\mathfrak{R})$, which is our goal). The reason is that here it is a priori possible that V may be *reducible*, i.e. it may consist of several components, say of equal dimension, one or more (but not all) of which constitute T_4. Then T_4 is still a proper algebraic subvariety of V, but in no unbiased sense does it have measure zero in V. This problem did not arise in the proof of (ii), for there we were showing that T_2 has measure zero in T_0, and $T_0=\mathbb{C}^{12}$ is *irreducible*.

The approach which we take here will avoid confrontation with the question of the irreducibility of V itself; we will work directly with $V(\mathfrak{R})$, which is of course our real interest for the purposes of this paper. The first step is to prove the following:

(3.7) *There exists a measure zero subvariety D of $V(\mathfrak{R})$ and a smooth, connected algebraic manifold M which admits a surjective algebraic map α onto $V(\mathfrak{R})-D$*

$\alpha: M \to V(\mathfrak{R})-D$.

Before we give the proof of (3.7), we will show how it leads to the desired result. The argument is a generalization of that of the last paragraph of the proof of (3.5) (ii) above. Assume that $T_4(\mathfrak{R})$ has measure greater than zero in $V(\mathfrak{R})$. Then there is some non-empty open set U of $V(\mathfrak{R})$ in which $T_4(\mathfrak{R})$ is dense. Since D has measure zero in $V(\mathfrak{R})$, we may assume U does not meet D. On the other hand, by upper semi continuity T_4 is an algebraic subvariety of V, and is a *proper* subvariety because of the point $P \in V(\mathfrak{R})-T_4(\mathfrak{R})$ which we will produce below. We will also show that the point P is not in D. Therefore T_4 is contained in the zero set of a polynomial f on $\mathbb{C}^{12}$ with complex coefficients, and this polynomial is not uniformly zero on $V(\mathfrak{R})-D$.

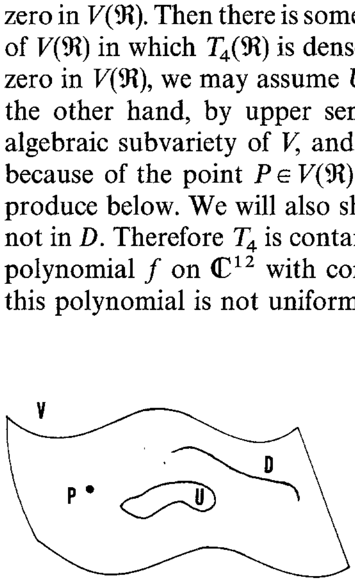

Fig. 3. Geometry underlying the proof of uniqueness using upper semi continuity

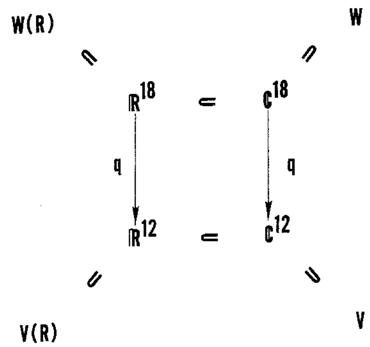

Fig. 4. The relationships among $W(\mathfrak{R})$, $V(\mathfrak{R})$, W, and V

As above, if we restrict to $\mathfrak{R}^{12}$ we can write the complex polynomial f in the form $g+\sqrt{-1}h$, where g and h are real polynomials not both of which are identically zero. For f to vanish at a point in $V(\mathfrak{R})$ (or indeed at any point of $\mathfrak{R}^{12}$), both g and h must vanish there. Thus the zero sets of the real polynomials g and h must be dense in U so that both g and h vanish identically on U.

It follows that the zero sets of the composite functions $g \circ \alpha$ and $f \circ \alpha$ vanish in the non-empty open set $\alpha^{-1}(U)$. Since g, h, and α are algebraic maps they are a fortiori real-analytic, so that $g \circ \alpha$ and $h \circ \alpha$ are real analytic functions on M. Since M is smooth and connected, $g \circ \alpha$ and $h \circ \alpha$ must therefore be identically zero on M. Hence g and h are identically zero on $V(\mathfrak{R})-D$, so also f is identically zero on $V(\mathfrak{R})-D$, a contradiction.

We now turn to the proof of (3.7). Recall that the set of points $W(\mathfrak{R})$ of W which have real coordinates may be interpreted as the set of all possible choices of three positions which can be occupied by a rigid configuration consisting of two arbitrary points and the origin, as it rotates about some fixed axis through the origin in $\mathfrak{R}^3$. In fact the eighteen coordinates of the point of $W(\mathfrak{R})$ are the (x, y, z) coordinates of the three successive positions of each of the two points. As before we will let $q: \mathbb{C}^{18} \to \mathbb{C}^{12}$ be the projection onto the X and Y components, i.e. $q(\{X_{ij}, Y_{ij}, Z_{ij}\})=\{(x_{ij}, y_{ij})\}$. It is clear that $q(W)=V$, and $q(W(\mathfrak{R}))=V(\mathfrak{R})$.

The eighteen coordinates of a point A in $W(\mathfrak{R})$ are the coordinates in $\mathfrak{R}^3$ of the points $A_{11}, A_{21}, A_{12}, A_{22}, A_{13}, A_{23}$, where A_{ij} denotes the position of A_i $(i=1,2)$ in each of the three views $(j=1,2,3)$. We will make the identification $A=(A_{11}, A_{12}, \ldots, A_{13}, A_{23})$. Let S denote the subset of $\mathrm{SO}(3,R) \times \mathrm{SO}(3,R)$ consisting of those pairs (σ, τ) such that $\sigma\tau=\tau\sigma$. It is well known that σ, τ commute if and only if they are rotations about the same axis in $\mathfrak{R}^3$ (which includes the case where either one of them is the identity). Thus the set S corresponds to all possible sequences of two rigid motions of $\mathfrak{R}^3$ which can be interpreted as successive rotations about the same axis; one or both of these rotations may be trivial, corresponding to the cases where one or both of σ, τ are the identity.

We can then see that there is a natural map

$$\pi: \mathfrak{R}^3 \times \mathfrak{R}^3 \times S \to W(\mathfrak{R})$$

defined by $\pi(v_1, v_2, \sigma, \tau) = (v_1, v_2, \sigma v_1, \sigma v_2, \tau v_1, \tau v_2)$, i.e. $\pi(v_1, v_2, \sigma, \tau)$ is the point A of $W(\mathfrak{R})$ with $A_{11} = v_1$, $A_{21} = v_2$, $A_{12} = \sigma v_1$, $A_{22} = \sigma v_2$, $A_{13} = \tau v_1$, $A_{23} = \tau v_2$. π is continuous since nearby rotations yield nearby points, and is surjective in view of the geometric description of $W(\mathfrak{R})$. Note moreover that the map π is algebraic, since $\pi(v_1, v_2, \sigma, \tau)$ can be computed explicitly in terms of polynomials in the coordinates of v_1, v_2 and the entries of the matrices which represent σ and τ. Since $q: W(\mathfrak{R}) \to V(\mathfrak{R})$ is also algebraic and surjective, we get:

$$q \circ \pi: \mathfrak{R}^3 \times \mathfrak{R}^3 \times S \to V(\mathfrak{R})$$

is a surjective algebraic map.

The dimension of S is four, since we have σ varying in $SO(3, R)$ whose dimension is three, and for each σ (other than the identity) we have one dimension of freedom for τ, i.e. an angle of rotation about the same axis as σ. When σ = identity, τ can be any element of $SO(3, R)$, but this contributes a three-dimensional subspace of S, so the overall dimension of four is unaffected. The dimension of $\mathfrak{R}^3 \times \mathfrak{R}^3 \times S$ is therefore ten.

Let E denote the set of points (v_1, v_2) in $\mathfrak{R}^3 \times \mathfrak{R}^3$ which are linearly dependent. E is a four-dimensional algebraic variety in $\mathfrak{R}^3 \times \mathfrak{R}^3$ and in particular $E \times S$ has dimension eight in $\mathfrak{R}^3 \times \mathfrak{R}^3 \times S$. Notice that outside of $E \times S$, π is an isomorphism onto $W(\mathfrak{R}) - \pi(E \times S)$. In fact, if v_1, v_2 are linearly independent, any rotations σ, τ are uniquely determined by their effect on v_1 and v_2. Therefore the dimension of $W(\mathfrak{R})$ is also ten, since the dimension of $\pi(E \times S)$ is at most eight, and π is surjective. The point is that $\pi: \mathfrak{R}^3 \times \mathfrak{R}^3 \times S \to W(\mathfrak{R})$ is a surjective algebraic map between algebraic varieties of the same dimension. Now $q: W(\mathfrak{R}) \to V(\mathfrak{R})$ is also surjective. Moreover there is at least one point P of $V(\mathfrak{R})$ for which $q^{-1}(P)$ is a finite set (for example the point P we will produce below). Hence, by a straightforward application of the upper semi continuity principle, there is a nonempty *open* set of $V(\mathfrak{R})$ over which q is finite-to-one. It follows that the dimension of $V(\mathfrak{R})$ is the same as $W(\mathfrak{R})$, i.e. $\dim V(\mathfrak{R}) = \dim W(\mathfrak{R}) = 10$. [Note it is a priori possible that $V(\mathfrak{R})$ has components of lower dimension.] Thus:

$$q \circ \pi: \mathfrak{R}^3 \times \mathfrak{R}^3 \times S \to V(\mathfrak{R})$$

is a surjective algebraic map between algebraic varieties of the same dimension.

Now let $D_1' = \{(\sigma, \tau) \in S \mid \text{either } \sigma \text{ or } \tau \text{ is the identity}\}$ and let $D' = \mathfrak{R}^3 \times \mathfrak{R}^3 \times D_1'$. D_1' is three-dimensional; it has two three-dimensional components $SO(3, \mathfrak{R}) \times \{\text{identity}\}$ and $\{\text{identity}\} \times SO(3, \mathfrak{R})$. Hence the dimension of D' is nine. Let $D = q \circ \pi(D')$. The dimension of D is at most nine, so since $\dim(V(\mathfrak{R}))$ is ten *D has measure zero in $V(\mathfrak{R})$*. Let $M = \mathfrak{R}^3 \times \mathfrak{R}^3 \times S - D' = \mathfrak{R}^3 \times \mathfrak{R}^3 \times (S - D_1')$, and let α denote the restriction of $q \circ \pi$ to M, $\alpha: M \to V(\mathfrak{R}) - D$. α is surjective and algebraic, and D has measure zero in $V(\mathfrak{R})$. To satisfy the hypothesis of (3.7), it remains to show that M is a smooth, connected manifold.

For this, it is obviously sufficient to show that $S - D_1'$ is a smooth connected manifold. Consider the map $p: (S - D_1') \to SO(3, R) - \{\text{identity}\}$, defined by $p(\sigma, \tau) = \sigma$. p is surjective, and $SO(3, R) - \{\text{identity}\}$ is a smooth, connected manifold. For any $\sigma \in SO(3, R) - \{\text{identity}\}$, $p^{-1}(\sigma)$ may be identified with the open interval $(0, 2\pi)$ in R, i.e. the set of all nontrivial rotations about the same axis as σ. It follows that $S - D_1'$ is a fibre bundle over $SO(3, R)$ with fibre $(0, 2\pi)$, so it is also a smooth connected manifold.

This almost concludes the proof of (3.7); the final order of business is to produce the point P in $V(\mathfrak{R}) - T_4$. For this consider $P \in \mathfrak{R}^{12}$ with coordinates:

$$(x_{11}, y_{11}) = (2.71076, 2.57115),$$
$$(x_{12}, y_{12}) = (2.57398, 1.99999),$$
$$(x_{13}, y_{13}) = (2.47320, 1.36808),$$
$$(x_{21}, y_{21}) = (5.48447, -1.92836),$$
$$(x_{22}, y_{22}) = (5.58706, -1.49999),$$
$$(x_{23}, y_{23}) = (5.66265, -1.02606).$$

We note immediately that $P \notin D$, for otherwise, by definition of D, either the second or third line in the above list of coordinates would be equal to the first line. Next one checks that the resulting system (3.6) has no solutions at infinity, i.e. when $W = 0$. Next, we can verify that the only solutions to (3.4) are $(z_{11}, z_{21}, z_{12}, z_{22}, z_{13}, z_{23}) = (-4.2473, 0.44941, -4.6231, 0.73127, -4.9000, 0.93895)$ and $(z_{11}, z_{21}, z_{12}, z_{22}, z_{13}, z_{23}) = (4.2473, -0.44941, 4.6231, -0.73127, 4.9000, -0.93895)$. One way to do this is first to generate the 64 solutions to the first six equations of (3.4) by the method of Sect. 2, and then test each of these solutions on the equations (3.4g and h). Thus $P \notin T_4$, provided that the solutions $\pm(z_{11}, \ldots, z_{23})$ above have *multiplicity one*. To eliminate the possibility of higher multiplicities, observe that any solution of multiplicity greater than one corresponds to a singular point of the solution set to (2.3) and (2.4) (for the given values of the parameters). Thus a multiple solution $(z_{11}, \ldots, z_{23})$ corresponds to a degeneracy in the Jacobian matrix (with respect to the z-coordinates) of these six equations. This matrix is

$$J=\begin{pmatrix} 2z_{11} & -2z_{12} & 0 & 0 & 0 & 0 \\ 2z_{11} & 0 & -2z_{13} & 0 & 0 & 0 \\ 0 & 0 & 0 & 2z_{21} & -2z_{22} & 0 \\ 0 & 0 & 0 & 2z_{21} & 0 & -2z_{23} \\ z_{21} & -z_{22} & 0 & z_{11} & -z_{12} & 0 \\ z_{21} & 0 & -z_{23} & z_{11} & 0 & -z_{13} \end{pmatrix}$$

which has determinant:

$$\det(J)=(z_{11}z_{22}-z_{12}z_{21})(z_{11}z_{23}-z_{13}z_{21}) \cdot (z_{12}z_{23}-z_{13}z_{22}).$$

We simply observe that the solutions corresponding to our point P above yield a non-zero value for $\det(J)$. This concludes the proof of (3.7), and hence of our Theorem 3.0.

Remark. The reader may ask whether it would not be better to *compute explicitly* the locus $T_4(\mathfrak{R})$, and observe directly its geometry in $V(\mathfrak{R})$. The answer is certainly yes, provided that the computation can be effectively carried out in a manner which yields a geometrically interpretable result. While one has the feeling that this is possible, we have been unable to find such a direct approach which would be feasible for presentation.

4 Induced Fixed-Axis Motion

The analysis of fixed-axis motion in Sect. 3 assumes that the axis of rotation is in a generic orientation with respect to the observer: the axis is neither parallel to the observer's line of sight nor is it perpendicular. From a purely mathematical point of view this would seem a quite weak assumption since the probability of these special orientations is zero. However, because one often translates along straight paths in environments that are largely static, one frequently observes fixed-axis motion where the axis lies orthogonal to the line of sight. The position of the axis depends upon which point of the visible environment one foveates.

In this section we investigate briefly the fixed-axis motion induced by a translating observer, showing that the axis of rotation is indeed orthogonal to the line of sight and giving a simple expression for the angular velocity induced by straight line translations. In Sect. 5 we consider the recovery of three-dimensional structure from fixed-axis motion in this special case, with the added restriction that the induced angular velocity is constant. We conclude that only three views of two points are needed. In Sect. 6 we eliminate the angular velocity constraint and provide a closed form solution for the three-dimensional structure given three views of three points.

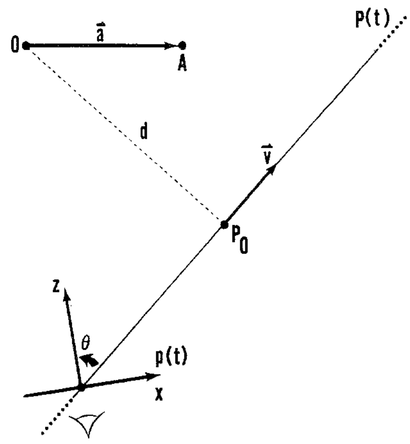

Fig. 5. Geometry underlying the derivation of the fixed-axis motion induced by a translating observer who foveates a fixed point

Consider an observer traveling along a straight path, $\mathbf{P}(t)$, given in $\mathfrak{R}^3$ by $\mathbf{P}(t)=\mathbf{P}_0+t\mathbf{v}$. The observer foveates some point O as he translates. Erect a coordinate system that translates with the observer such that the plane defined by $\mathbf{P}(t)$ and O is the xz-plane. Figure 3 gives a top view of this plane. Further, choose the coordinate system so that the effect of foveating O is to make O's x coordinate zero (O's y coordinate is also zero since O lies in the xz-plane.) Consider the effect of the observer's translation on some vector $\mathbf{a}$ from O to some point A. The effect is simply to translate the tail of the vector along the z-axis of the observer's moving coordinate system and then to rotate the vector about the y-axis. Any translation of the vector parallel to the observer's image plane is nullified by foveation. And, assuming orthographic viewing, any translation of the vector along the z-axis has no visible effect. The net result is that the vector undergoes a rotation about a fixed axis, in fact about the y-axis, which is orthogonal to the observer's line of sight (the z-axis). This holds true for vectors that lie in the xz-plane, as shown in the figure, as well as for those that do not.

Let $\mathbf{P}_0$ be the point along $\mathbf{P}(t)$ of minimum distance from the foveated point. Then the induced angular velocity at time t depends upon the magnitude of $\mathbf{P}(t)-\mathbf{P}_0$ (say $p(t)$), the observer's velocity $p'(t)$, and the minimum distance, d, from his path to the foveated point. From Fig. 5 one sees that the angle between the observer's z-axis and $\mathbf{P}(t)$ is $\theta(t)=\tan^{-1}(d/p(t))$. The change in this angle is precisely the amount that the vector $\mathbf{a}$ rotates. Thus the induced angular velocity is

$$\theta'(t)=\frac{-dp'(t)}{d^2+p^2(t)} \tag{4.1}$$

and the induced angular acceleration is

$$\theta''(t) = \frac{2dp(t)(p'(t))^2}{(d^2+p^2(t))^2} - \frac{dp''(t)}{d^2+p^2(t)}, \quad (4.2)$$

which is zero only if

$$p''(t) = \frac{2p(t)(p'(t))^2}{d^2+p^2(t)}. \quad (4.3)$$

If the observer travels at a constant velocity, say $p'(t)=c$, then the induced angular acceleration is not zero, but is

$$\theta''(t) = \frac{2dc^2p(t)}{(d^2+p^2(t))^2}. \quad (4.4)$$

5 Fixed-Axis Motion: No Angular Acceleration

In this section we prove the following:

Theorem 5.0 *Given three orthographic projections of two points spinning rigidly and at a constant angular velocity about a fixed axis that is parallel to the image plane, the three-dimensional structure and motion of the points is uniquely determined up to a reflection about the image plane.*

As discussed in the previous section, the motivation for examining this special case is that the axis of rotation induced by a translating observer is orthogonal to his line of sight. However, as indicated by (4.2), the induced angular velocity is not likely to be constant. We consider the case of constant angular velocity anyway because it leads to a particularly simple solution and because the induced angular acceleration is small when the observer is distant from his point of nearest approach to the foveated point.

We assume, without loss of generality, that the observer is foveating one of the two points and that the successive positions of the other point over the three views lies on a line parallel to the x-axis of the observer's coordinate system. A top view and a side view of this geometry are shown in Fig. 6. Let x_j be the x coordinate of the point in view j, where $j=1,2,3$. Let r be the radius of the circular path traced out by the point. Let θ be the angle between the image plane and the vector from the origin to the point in the first view. Let δ be the (constant) angular rotation between views. Then we can write:

$$x_1 = r\cos(\theta),$$
$$x_2 = r\cos(\theta+\delta), \quad (5.1)$$
$$x_3 = r\cos(\theta+2\delta).$$

Using the double angle formulae for sines and cosines, (5.1) becomes:

$$x_1 = r\cos(\theta), \quad (5.2a)$$
$$x_2 = r[\cos(\theta)\cos(\delta) - \sin(\theta)\sin(\delta)], \quad (5.2b)$$
$$x_3 = r[\cos(\theta)\cos^2(\delta) - \cos(\theta)\sin^2(\delta) - 2\sin(\theta)\sin(\delta)\cos(\delta)]. \quad (5.2c)$$

Dividing (5.2a) into (5.2b) and (5.2c) gives the two equations

$$\frac{x_2}{x_1} = \cos(\delta) - \tan(\theta)\sin(\delta), \quad (5.3a)$$

$$\frac{x_3}{x_1} = \cos^2(\delta) - \sin^2(\delta) - 2\tan(\theta)\sin(\delta)\cos(\delta). \quad (5.3b)$$

Equation (5.3a) can be solved for $\tan(\theta)$,

$$\tan(\theta) = \frac{\cos(\delta) - x_2/x_1}{\sin(\delta)}, \quad (5.4)$$

and substituted into (5.3b) to give

$$\frac{x_3}{x_1} = \frac{2x_2}{x_1}\cos(\delta) - \sin^2(\delta) - \cos^2(\delta). \quad (5.5)$$

Solving (5.5) for $\cos(\delta)$ gives

$$\cos(\delta) = \frac{x_3+x_1}{2x_2}. \quad (5.6)$$

Once δ is known from (5.6), one can determine θ from (5.4) and finally r from (5.2a). Consequently the three-

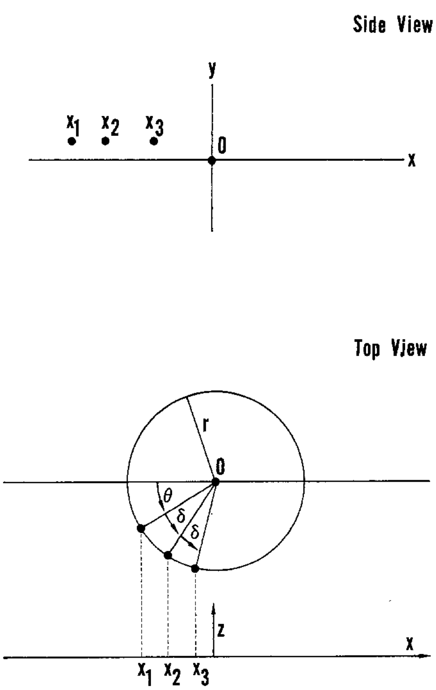

Fig. 6. Geometry underlying the computation of structure from three orthographic views of two points that spin at a constant angular velocity about an axis parallel to the image plane

dimensional interpretation is unique up to a reflection. The reflective ambiguity arises from (5.6) because knowing the cosine of an angle only specifies the angle up to a sign.

The probability of false targets in this analysis is the probability that six randomly chosen points lie on two parallel lines – which is zero.

6 Fixed-Axis Motion: Angular Acceleration

In this section we prove the following:

Theorem 6.0. *Given three orthographic views of three points spinning rigidly about a fixed axis that is parallel to the image plane, there are at most two interpretations (plus reflections) for the three-dimensional structure and motion of the points. In particular, constant angular velocity need not be assumed.*

The geometry for this proof is shown in Fig. 7. We again assume, without loss of generality, that one of the three points is foveated by the observer and that the other points move along lines parallel to the x-axis of the observer's coordinate system. Let x_{ij} be the x coordinate of point i in view j, where $i=1,2$ and $j=1,2,3$. Let θ_j be the angle between the image plane and the vector from the origin to the first point in view j. Let β_i be the angle made by the vector from the origin to the first point with the vector from the origin to point i. (Note that $\beta_1=0$). Finally, let r_i be the radius of the circular path traced out by point i. Then we can write the six equations

$$x_{ij}=r_i\cos(\theta_j+\beta_i), \quad i=1,2; \quad j=1,2,3. \tag{6.1}$$

Let

$$\begin{aligned} s_i &= 1/r_i, \\ z_j &= \cos(\theta_j), \\ w_j &= \sin(\theta_j), \\ u_i &= \cos(\beta_i), \\ v_i &= \sin(\beta_i). \end{aligned} \tag{6.2}$$

Then from the first view we have (using the double-angle formula for cosines)

$$x_{11}s_1=z_1, \tag{6.3a}$$

$$x_{21}s_2=z_1u_2-w_1v_2. \tag{6.3b}$$

From the second view we have

$$x_{12}s_1=z_2, \tag{6.4a}$$

$$x_{22}s_2=z_2u_2-w_2v_2. \tag{6.4b}$$

From the third view we have

$$x_{13}s_1=z_3, \tag{6.5a}$$

$$x_{23}s_2=z_3u_2-w_3v_2. \tag{6.5b}$$

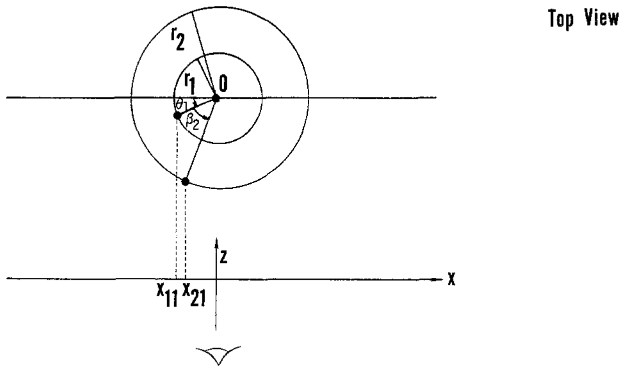

Fig. 7. Geometry underlying the computation of structure from three orthographic views of three points that spin with arbitrary angular accelerations about an axis parallel to the image plane

Dividing (6.4a) by (6.3a), and (6.4b) by (6.3b), gives, respectively,

$$\frac{x_{12}}{x_{11}}z_1=z_2, \tag{6.6a}$$

$$\frac{x_{22}}{x_{21}}(z_1u_2-w_1v_2)=(z_2u_2-w_2v_2). \tag{6.6b}$$

Dividing (6.5a), by (6.3a), and (6.5b) by (6.3b) gives, respectively,

$$\frac{x_{13}}{x_{11}}z_1=z_3, \tag{6.7a}$$

$$\frac{x_{23}}{x_{21}}(z_1u_2-w_1v_2)=(z_3u_2-w_3v_2). \tag{6.7b}$$

Eliminate z_2 from (6.6b) using (6.6a):

$$\left(\frac{x_{22}}{x_{21}}-\frac{x_{12}}{x_{11}}\right)z_1u_2=\left(\frac{x_{22}}{x_{21}}w_1-w_2\right)v_2. \tag{6.8}$$

Eliminate z_3 from (6.7b) using (6.7a):

$$\left(\frac{x_{23}}{x_{21}}-\frac{x_{13}}{x_{11}}\right)z_1u_2=\left(\frac{x_{23}}{x_{21}}w_1-w_3\right)v_2. \tag{6.9}$$

Multiply (6.8) by $x_{23}/x_{21}-x_{13}/x_{11}$; multiply (6.9) by $x_{22}/x_{21}-x_{12}/x_{11}$. Subtract (6.9) from (6.8) and simplify:

$$\begin{aligned}(x_{12}x_{23}-x_{13}x_{22})w_1+(x_{13}x_{21}-x_{23}x_{11})w_2 \\ +(x_{22}x_{11}-x_{21}x_{12})w_3=0.\end{aligned} \tag{6.10}$$

(Interestingly, this can be written as $[(x_{11}, x_{12}, x_{13}) \times (x_{21}, x_{22}, x_{23})]\cdot(w_1, w_2, w_3)=0$.) Recalling that $z_i^2+w_i^2=1$, we can rewrite (6.6a) and (6.7a) as

$$\left(\frac{x_{12}}{x_{11}}\right)^2(1-w_1^2)=1-w_2^2, \tag{6.11a}$$

$$\left(\frac{x_{13}}{x_{11}}\right)^2(1-w_1^2)=1-w_3^2. \tag{6.11b}$$

(6.10) and (6.11) give us three equations in three of the unknowns w_1, w_2, w_3. We will use them to derive a closed form solution for these unknowns.

Multiply (6.11a) by x_{13}^2/x_{11}^2-1. Multiply (6.11b) by x_{12}^2/x_{11}^2-1. Subtract (6.11b) from (6.11a) and simplify:

$$(x_{13}^2-x_{12}^2)w_1^2+(x_{11}^2-x_{13}^2)w_2^2 +(x_{12}^2-x_{11}^2)w_3^2=0. \tag{6.12}$$

Solve (6.10) for w_1 to get

$$w_1=-a^{-1}(bw_2+cw_3), \tag{6.13}$$

where

$$a=x_{23}x_{12}-x_{22}x_{13}, \tag{6.14a}$$

$$b=x_{13}x_{21}-x_{23}x_{11}, \tag{6.14b}$$

$$c=x_{22}x_{11}-x_{12}x_{21}. \tag{6.14c}$$

Substitute (6.13) into (6.12) to give

$$(x_{13}^2-x_{12}^2)(bw_2+cw_3)^2+a^2(x_{11}^2-x_{13}^2)w_2^2 +a^2(x_{12}^2-x_{11}^2)w_3^2=0, \tag{6.15}$$

which may be simplified further to

$$\alpha w_3^2+\beta w_2w_3+\gamma w_2^2=0, \tag{6.16}$$

where

$$\alpha=a^2(x_{12}^2-x_{11}^2)+c^2(x_{13}^2-x_{12}^2), \tag{6.17a}$$

$$\beta=2bc(x_{13}^2-x_{12}^2), \tag{6.17b}$$

$$\gamma=a^2(x_{11}^2-x_{13}^2)+b^2(x_{13}^2-x_{12}^2). \tag{6.17c}$$

Divide (6.16) by w_2^2 and solve for w_3/w_2 using the quadratic formula:

$$\zeta=\frac{w_3}{w_2}=\frac{-\beta\pm\sqrt{\beta^2-4\alpha\gamma}}{2\alpha}. \tag{6.18}$$

Having a value for the ratio of w_3 and w_2, we can return to (6.11) to get w_2 and w_3 explicitly. Multiply (6.11a) by x_{13}^2/x_{11}^2. Multiply (6.11b) by x_{12}^2/x_{11}^2. Subtract (6.11b) from (6.11a):

$$x_{12}^2w_3^2-x_{13}^2w_2^2+x_{13}^2-x_{12}^2=0. \tag{6.19}$$

Equation (6.19) can be reexpressed in terms of ζ, (i.e., w_3/w_2),

$$x_{12}^2\zeta^2w_2^2-x_{13}^2w_2^2+x_{13}^2-x_{12}^2=0, \tag{6.20}$$

and solved for w_2:

$$w_2=\pm\sqrt{\frac{x_{12}^2-x_{13}^2}{x_{12}^2\zeta^2-x_{13}^2}}. \tag{6.21}$$

Having w_2, we can solve for w_3 using (6.19), and then solve for w_1 using (6.10). Then using the fact that $z_i^2+w_i^2=1$, we can find z_1, z_2, and z_3.

To find u_2 and v_2, multiply (6.3b) by x_{22} and (6.4b) by x_{21}. Subtract (6.4b) from (6.3b), and solve for u_2 in terms of v_2:

$$u_2=v_2\left(\frac{x_{22}w_1-x_{21}w_2}{x_{22}z_1-x_{21}z_2}\right). \tag{6.22}$$

Use the fact that $u_2^2+v_2^2=1$ to solve for u_2 and v_2:

$$v_2=\pm\left(\left(\frac{x_{22}w_1-x_{21}w_2}{x_{22}z_1-x_{21}z_2}\right)^2+1\right)^{-1/2}, \tag{6.23a}$$

$$u_2=\pm(1-v_2^2)^{-1/2}. \tag{6.23b}$$

Finally, from (6.3) we can find the radii of the circular paths, r_1 and r_2.

The probability of false targets in this analysis is the probability that nine randomly chosen points in the plane lie on three parallel straight lines – which is zero.

7 Conclusion

The principal results discussed in this paper are the following. Three views of three points in a rigid configuration lead to two interpretations of the three-dimensional structure (plus orthographic reflections). Each of these four structures has sixteen possible motions, leading to a total of sixty four interpretations of structure and motion. Adding a fourth point, as Ullman (1979) has shown, leads to a unique interpretation. Assuming fixed-axis motion also leads to a unique interpretation for three views of three points if the axis is in a generic orientation. If the axis is parallel to the image plane then three views of the three points are compatible with at most two interpretations – unless one assumes the angular velocity is constant, in which case only two points are needed and one obtains a unique interpretation. Closed form solutions are obtained for each result.

The equations studied here are amenable to solution by the techniques of nonlinear programming, making possible the design of noise insensitive algorithms for machine vision systems. The closed form solutions presented in the paper are, of course, unsuitable as machine vision algorithms – they are presented only to prove that in fact the equations have a unique solution. However the equations themselves can be combined into an objective function which is minimized using any of several nonlinear optimization techniques. An example of this is given by Reuman and Hoffman (1986), who devise noise insensitive algorithms for the equations studied by Hoffman and Flinchbaugh (1982).

It may seem natural to ask whether it is possible for the human visual system to employ the type of processor described in this paper, and in conclusion we will briefly address this question. The first question is whether humans are capable of detecting a structure rotating rigidly about a fixed axis given three views of three points on it. Let us assume the answer is affirmative (as the evidence from Braunstein's pilot studies indicates). This *means* that the visual system computes the variety $W(\mathfrak{R})$ given $V(\mathfrak{R})$; in fact the perception of the structure consists in knowing the z-coordinates given the x- and y-coordinates, and this is *exactly* the information encoded by the varieties W and V and the projection from W to V. Secondly, since V and W are algebraic varieties, knowledge of them is exactly equivalent to knowledge of the set of polynomials which vanish on them (i.e. of their largest "ideal" of definition). In our case it is not hard to show that this set of polynomials is generated irredundantly by our equations (2.3), (2.4), and (3.2). The point is then that the ability to perceive fixed-axis motion from three views of three points is informationally equivalent to the solution of these equations, or of some equivalent set of equations related to these by a change of coordinates. This is true a priori (i.e. this truth is algorithm independent). One may ask for example about the *way* in which these equations are solved in some system capable of this type of perception. But the *fact* that the equations are solved is precisely equivalent to the capability. From this point of view we can see that the most natural rhetorical question is whether it is possible for the human visual system *not* to employ the processor discussed above. And we are suggesting that the natural answer to this question is no.

Acknowledgements. We thank G.J. Andersen, M. Braunstein, W. Richards, and R. Yoshino for useful discussions. This material is based upon work supported by the National Science Foundation under Grant No. IST-8413560, and by a contract to D. Hoffman from the Office of Naval Research, Psychological Sciences Division, Engineering Psychology Group.

References

Atiyah-Macdonald M (1969) Introduction to commutative algebra. Addison-Wesley, Reading, MA

Bennett BM, Hoffman DD (1985) The computation of structure from fixed-axis motion: nonrigid structures. Biol Cybern 51:293–300

Bobick A (1983) A hybrid approach to structure-from-motion. Proceedings of the ACM Workshop on Motion: Representation and perception

Braunstein M (1976) Depth perception through motion. Academic Press, London

Gibson JJ (1950) The perception of the visual world. Houghton Mifflin, Boston

Gibson JJ, Gibson EJ (1957) Continuous perspective transformations and the perception of rigid motion. J Exp Psychol 54:129–138

Green BF (1961) Figure coherence in the kinetic depth effect. J Exp Psychol 62:272–282

Grothendieck A (1961) Eléments de géometrie algebrique. III. Etude cohomologique des faisceaux cohérents. Publications Mathematiques de l'Institut des Hautes Etudes Scientifiques 11

Hartshorne R (1977) Algebraic geometry. Springer, New York Berlin Heidelberg

Hay CJ (1966) Optical motions and space perception – an extension of Gibson's analysis. Psychol Rev 73:550–565

Helmholtz H (1925) Treatise one physiological optics. In: Southall JPC (ed) Dover, New York

Hoffman DD (1982) Inferring local surface orientation from motion fields. J Opt Soc Am 72:888–892

Hoffman DD (1983) The interpretation of visual illusions. Sci Am 249:154–162

Hoffman DD, Flinchbaugh BE (1982) The interpretation of biological motion. Biol Cybern 42:195–204

Johansson G (1975) Visual motion perception. Sci Am 232:76–88

Longuet-Higgins HC, Prazdny K (1980) The interpretation of a moving retinal image. Proc R Soc Lond Ser B 208:385–397

Reuman S, Hoffman DD (1986) Regularities of nature: their role in the interpretation of visual motion. In: Pentland AP (ed) From pixels to predicates: recent advances in computational vision. Ablex Publishing, New York (in press)

Ullman S (1979) The interpretation of structure from motion. Proc R Soc Lond Ser B 203:405–426

Wallach H, O'Connell D (1953) The kinetic depth effect. J Exp Psychol 45:205–217

Webb J, Aggarwal J (1982) Structure from motion of rigid and jointed objects. Artif Intell 19:107–130

Received: June 3, 1985

Prof. Donald D. Hoffman
School of Social Sciences
University of California, Irvine
Irvine, CA 92717
USA