J-RA/4/2//18919

A Simple Procedure to Solve Motion and Structure from Three Orthographic Views

Xinhua Zhuang
Thomas S. Huang
Robert M. Haralick

Reprinted from
IEEE JOURNAL OF ROBOTICS AND AUTOMATION
Vol. 4, No. 2, April 1988

A Simple Procedure to Solve Motion and Structure from Three Orthographic Views

XINHUA ZHUANG, THOMAS S. HUANG, AND ROBERT M. HARALICK

Abstract—**In his influential work on motion analysis [1], Ullman showed that for the orthographic case, 4-point correspondences over three views are sufficient to determine motion and structure of the 4-point rigid configuration. However, his method is nonlinear and the conditions of uniqueness and convergence of his algorithm are not made clear. In this communication, we show a very simple linear algorithm to solve Ullman's problem of determining motion and structure from three orthographic views. We also give necessary and sufficient conditions of a unique solution. This research paves the way towards a robust version of the algorithm which, as a viable and practical means for computing structure and motion, may have a potential for wide applications.**

Manuscript received November 20, 1986; revised March 10, 1987.
X. Zhuang and T. S. Huang are with the Coordinated Science Laboratory, University of Illinois, Urbana, IL 61801.
R. M. Haralick is with the Department of Electrical Engineering, University of Washington, Seattle, WA 98195.
IEEE Log Number 8718918.

0882-4967/88/0400-0236$01.00

I. Introduction

As is well known, the topic of the motion and surface structure estimates from multiple views (perspective projections or orthographic projections) is important to the field of Computational Vision. In his influential work on motion analysis [1], Ullman showed that for the orthographic case, 4-point correspondences over three views are sufficient to determine motion and structure of the 4-point rigid configuration. However, his method is nonlinear and conditions of uniqueness and convergence of his algorithm are not clear. The problem has been re-examined recently by Aloimonos and Brown [2] and Lee [3]. Noticeably, Lee's technique is linear using only vector-matrix equations and vector cross products. An even simpler linear algorithm is possible. This communication will show that algorithm and give necessary and sufficient conditions of a unique solution. Doing so, we develop a useful formula (see (33)) as a by-product. In terms of this formula we can explain why, given two orthographic views, the motion and structure of a rigid object cannot be determined no matter how many point correspondences are used. This research work paves the way towards a robust version of the algorithm which, as a viable and practical means for computing structure and motion, may have a potential for wide applications since the algorithm requires fewer point correspondences than the well-known two views-motion algorithms (see Longuet-Higgins [4], Tsai and Huang [5], Zhuang *et al.* [6]) and since in many practical situations orthographic views give good approximation to the imaging processes.

Section II contains the problem formulation. In Section III we show how to uniquely determine r_{33}, s_{33}, $\pm(r_{13}, r_{23}, r_{31}, r_{32})$, $\pm(s_{13}, s_{23}, s_{31}, s_{32})$ given three orthographic views. In Section IV we show how to uniquely determine an orthonormal matrix R given $r_{13}, r_{23}, r_{31}, r_{32}$, r_{33} with $0 < r_{13}^2 + r_{23}^2 = r_{31}^2 + r_{32}^2 = 1 - r_{33}^2 \leqslant 1$. In Section V we give the simple procedure to uniquely determine the motion and structure from three orthographic views. The final section is a summary.

II. Problem Formulation

We assume that the image plane is stationary and that three orthographic views at time t_1, t_2, and t_3, respectively, are taken of a rigid object moving in the three-dimensional (3-D) object space. By processing the three views, we intend to determine the motion and structure of the 3-D object.

We shall use the following notations. Let (x, y, z) be the object space coordinates, and (X, Y) the image space coordinates. The X- and Y-axis coincide with the x- and y-axis (in particular, the origins of the x-y-z coordinate system and the X-Y coordinate system coincide). Let

(x, y, z)	object-space coordinates of a point P on the rigid object at t_1,
(x', y', z')	object-space coordinates of the same point P at t_2,
(x'', y'', z'')	object-space coordinates of the point P at t_3,
(X, Y)	image-space coordinates of the point P at t_1,
(X', Y')	image-space coordinates of the point P at t_2,
(X'', Y'')	image-space coordinates of the point P at t_3.

Then

$$(x', y', z')^t = R(x, y, z)^t + T_r \tag{1}$$

$$(x'', y'', z'')^t = S(x, y, z)^t + T_s \tag{2}$$

where $R \triangleq (r_{ij})_{3\times 3}$ and $S \triangleq (s_{ij})_{3\times 3}$ are rotation matrices, $T_r \triangleq (t_{ri})_{3\times 1}$ and $T_s \triangleq (t_{si})_{3\times 1}$ are 3×1 translation vectors. The superscript "t" represents transposition.

The problem we are trying to solve is: Given four image point correspondences

$$(X_i, Y_i) \leftrightarrow (X_i', Y_i') \leftrightarrow (X_i'', Y_i''), \qquad i=1, 2, 3, 4 \tag{3}$$

determine (R, T_r), (S, T_s), (x_i, y_i, z_i), $i = 1, 2, 3, 4$.

Note that with orthographic projections

$$(X, Y) = (x, y),$$
$$(X', Y') = (x', y'),$$
$$(X'', Y'') = (x'', y'') \tag{4}$$

and therefore it is obvious that t_{r3}, t_{s3} can never be determined and we can hope to determine the depths of the object points to only within an additive constant. What we are trying to determine are then: R, S, t_{ri}, $t_{si} (i = 1, 2)$, $z_i - z_1$, $i = 2, 3, 4$.

We can decompose the rigid body motion from t_1 to t_2 (t_3) as a rotation $R(S)$ around the point (x_1, y_1, z_1) followed by a translation $(x_1', y_1', z_1')^t$ ($(x_1'', y_1'', z_1'')^t$).

To determine R, S, $z_i - z_1$, $i = 2, 3, 4$, we can then use the following equations:

$$(X_i' - X_1', Y_i' - Y_1')^t = \bar{R}(X_i - X_1, Y_i - Y_1, z_i - z_1)^t$$
$$(X_i'' - X_1'', Y_i'' - Y_1'')^t = \bar{S}(X_i - X_1, Y_i - Y_1, z_i - z_1)^t$$
$$i = 2, 3, 4 \tag{5}$$

where $\bar{R} \triangleq (r_{ij})_{2\times 3}$, $\bar{S} \triangleq (s_{ij})_{2\times 3}$.

Once R, S, $z_i - z_1$, $i = 2, 3, 4$, all are determined, both $z_i' - z_1'$ and $z_i'' - z_1''$ $(i = 2, 3, 4)$ can be uniquely determined, and moreover both t_{ri} and t_{si} $(i = 1, 2)$ are determined as a function of z_1.

In the next section, we show that r_{33} and s_{33} can be uniquely determined, and $(r_{13}, r_{23}, r_{31}, r_{32})$ and $(s_{13}, s_{23}, s_{31}, s_{32})$ up to a sign can also be uniquely determined, both from three orthographic views.

III. Solving r_{33}, s_{33}, $(r_{13}, r_{23}, r_{31}, r_{32})$, $(s_{13}, s_{23}, s_{31}, s_{32})$

Let

$$a_1 = (X_2 - X_1, X_3 - X_1, X_4 - X_1)$$
$$a_2 = (Y_2 - Y_1, Y_3 - Y_1, Y_4 - Y_1)$$
$$a_3 = (z_2 - z_1, z_3 - z_1, z_4 - z_1)$$
$$b_1 = (X_2' - X_1', X_3' - X_1', X_4' - X_1')$$
$$b_2 = (Y_2' - Y_1', Y_3' - Y_1', Y_4' - Y_1')$$
$$c_1 = (X_2'' - X_1'', X_3'' - X_1'', X_4'' - X_1'')$$
$$c_2 = (X_2'' - X_1'', X_3'' - X_1'', X_4'' - X_1'')$$

$$A = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

$$\bar{A} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \qquad \bar{B} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \qquad \bar{C} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

Then $\bar{A}$, $\bar{B}$, $\bar{C}$ are known and (5) can be written as

$$\bar{B} = \bar{R}A. \tag{6}$$

$$\bar{C} = \bar{S}A. \tag{7}$$

We further assume

$$\text{Rank } [A] = 3. \tag{8}$$

Now what we are trying to determine are: R, S, a_3.

If

$$\text{Rank} \begin{bmatrix} \bar{A} \\ \bar{B} \end{bmatrix} < 3 \text{ (i.e., } = 2)$$

then $r_{13} = r_{23} = 0$ because of (8) and

$$\bar{B} = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} \bar{A} + \begin{bmatrix} r_{13} \\ r_{23} \end{bmatrix} a_3$$

and $r_{11}, r_{12}, r_{21}, r_{22}$ are uniquely determined by

$$\begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} = (\bar{B}\bar{A}^t)(\bar{A}\bar{A}^t)^{-1}. \tag{9}$$

It is thus obvious that $r_{33} = \pm 1$, $r_{31} = r_{32} = 0$, and r_{33} equals $+1$ or -1 depending on whether

$$\det \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix}$$

equals $+1$ or not because of

$$\det(R) = r_{33} \cdot \det \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix}.$$

In summary, when

$$\text{Rank} \begin{bmatrix} \bar{A} \\ \bar{B} \end{bmatrix} < 3$$

the rotation matrix R can be uniquely determined even though (6) does not contain any information about a_3.

If both

$$\text{Rank} \begin{bmatrix} \bar{A} \\ \bar{B} \end{bmatrix} \quad \text{and} \quad \text{Rank} \begin{bmatrix} \bar{A} \\ \bar{C} \end{bmatrix}$$

are less than 3, then both R and S can be uniquely determined. However, in this case the row vector a_3 cannot be recovered.

If

$$\text{Rank} \begin{bmatrix} \bar{A} \\ \bar{B} \end{bmatrix} < 3 \quad \text{and} \quad \text{Rank} \begin{bmatrix} \bar{A} \\ \bar{C} \end{bmatrix} = 3$$

then R is uniquely determined, and both S and a_3 have infinite solutions (see Section IV). Similarly, when

$$\text{Rank} \begin{bmatrix} \bar{A} \\ \bar{B} \end{bmatrix} = 3 \quad \text{and} \quad \text{Rank} \begin{bmatrix} \bar{A} \\ \bar{C} \end{bmatrix} < 3$$

S is uniquely determined, and both R and a_3 have infinite solutions. Therefore, the case which really interests us is

$$\text{Rank} \begin{bmatrix} \bar{A} \\ \bar{B} \end{bmatrix} = 3 \tag{10}$$

$$\text{Rank} \begin{bmatrix} \bar{A} \\ \bar{C} \end{bmatrix} = 3. \tag{11}$$

In the following we show how r_{33}, s_{33}, $(r_{13}, r_{23}, r_{31}, r_{32})$ up to a sign, $(s_{13}, s_{23}, s_{31}, s_{32})$ up to a sign can all be uniquely determined under the assumptions (10) and (11).

Since the rotation matrix R can be expressed as function of θ, ϕ, and ψ as follows:

$$R = \begin{bmatrix} \cos\psi\cos\theta & \sin\psi\cos\theta & -\sin\theta \\ -\sin\psi\cos\phi + \cos\psi\sin\theta\sin\phi & \cos\psi\cos\phi + \sin\psi\sin\theta\sin\phi & \cos\theta\sin\phi \\ \sin\psi\sin\phi + \cos\psi\sin\theta\cos\phi & -\cos\psi\sin\phi + \sin\psi\sin\theta\cos\phi & \cos\theta\cos\phi \end{bmatrix}$$

it is easy to directly verify the following identity:

$$(r_{32}, -r_{31}, 0) = (-r_{23}, r_{13})\bar{R}. \tag{12}$$

Similarly, it holds that

$$(s_{32}, -s_{31}, 0) = (-s_{23}, s_{13})\bar{S}. \tag{13}$$

Using (6), (7) and (12), (13), we obtain

$$(r_{32}, -r_{31}, 0)A = (-r_{23}, r_{13})\bar{B} \tag{14}$$

$$(s_{32}, -s_{31}, 0)A = (-s_{23}, s_{13})\bar{C} \tag{15}$$

or equivalently

$$(r_{32}, -r_{31})\bar{A} = (-r_{23}, r_{13})\bar{B} \tag{16}$$

$$(s_{32}, -s_{31})\bar{A} = (-s_{23}, s_{13})\bar{B} \tag{17}$$

or still equivalently

$$[\bar{A}^t \bar{B}^t](r_{32}, -r_{31}, r_{23}, -r_{13})^t = 0 \tag{18}$$

$$[\bar{A}^t \bar{C}^t](s_{32}, -s_{31}, s_{23}, -s_{13})^t = 0. \tag{19}$$

Because of the assumptions (10) and (11) both $(r_{13}, r_{23}, r_{31}, r_{32})$ up to a scale factor α and $(s_{13}, s_{23}, s_{31}, s_{32})$ up to a scale factor β are uniquely determined by (18) and (19). That is, there exists four 1×2 row vectors u, v, $\bar{u}$, $\bar{v}$ such that

$$\begin{aligned} (r_{13}, r_{23}) &= \alpha u \\ (r_{31}, r_{32}) &= \alpha v \\ (s_{13}, s_{23}) &= \beta \bar{u} \\ (s_{31}, s_{32}) &= \beta \bar{v}. \end{aligned} \tag{20}$$

Because of $r_{13}^2 + r_{23}^2 = r_{31}^2 + r_{32}^2 \leqslant 1$ and $r_{13}^2 + r_{23}^2 > 0$, the latter is implied by the assumption (10), we confirm

$$\alpha \neq 0 \quad \text{and} \quad \|u\| = \|v\| > 0. \tag{21}$$

Similarly, using the assumption (11) we confirm

$$\beta \neq 0 \quad \text{and} \quad \|\bar{u}\| = \|\bar{v}\| > 0. \tag{22}$$

Without loss of generality, in the proceeding discussion we assume

$$\|u\| = \|v\| = \|\bar{u}\| = \|\bar{v}\| = 1. \tag{23}$$

Using u, v, $\bar{u}$, $\bar{v}$ and (6), (7), we can further derive

$$\begin{aligned} u\bar{B} &= u \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} \bar{A} + u \begin{bmatrix} r_{13} \\ r_{23} \end{bmatrix} a_3 = u \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} \bar{A} + \alpha a_3 \\ &= (r_{11}r_{13} + r_{21}r_{23},\ r_{12}r_{13} + r_{22}r_{23})\bar{A} + \alpha a_3 \\ &= -r_{33}(r_{31}, r_{32})\bar{A} + \alpha a_3, \quad \text{since } r_{11}r_{13} + r_{21}r_{23} + r_{31}r_{33} = 0 \text{ and } r_{12}r_{13} + r_{22}r_{23} + r_{32}r_{33} = 0 \\ &= -r_{33} v\bar{A} + \alpha a_3 \end{aligned} \tag{24}$$

and

$$\bar{u}\bar{C} = -s_{33}\bar{v}\bar{A} + \beta a_3. \tag{25}$$

Equations (24) and (25) lead to

$$\beta[u\bar{B} + r_{33}(v\bar{A})] = \alpha[\bar{u}\bar{C} + s_{33}(\bar{v}\bar{A})] \tag{26}$$

which can be further written in a more readable form as

$$[(u\bar{B})^t, (v\bar{A})^t, (\bar{v}\bar{A})^t]\left(\frac{\beta}{\alpha}, \frac{\beta}{\alpha} r_{33}, -s_{33}\right)^t = (\bar{u}\bar{C})^t. \tag{27}$$

If the 3×3 coefficient matrix of (27) has a rank 3, then β/α, r_{33}, s_{33} are all uniquely determined, as can easily be seen. Furthermore, α and β are uniquely determined up to a sign

$$\alpha = \pm\sqrt{1-r_{33}^2} \tag{28}$$

$$\beta = \alpha \cdot \left(\frac{\beta}{\alpha}\right) = \pm\sqrt{1-r_{33}^2} \cdot \left(\frac{\beta}{\alpha}\right). \tag{29}$$

As a result, both $(r_{13}, r_{23}, r_{31}, r_{32})$ up to a sign and $(s_{13}, s_{23}, s_{31}, s_{32})$ up to a sign are uniquely determined

$$(r_{13} \quad r_{23} \quad r_{31} \quad r_{32}) = \pm\sqrt{1-r_{33}^2}(u, V) \tag{30}$$

$$(s_{13} \quad s_{23} \quad s_{31} \quad s_{32}) = \pm\left(\frac{\beta}{\alpha}\right)\sqrt{1-r_{33}^2}(\bar{u}, \bar{v}). \tag{31}$$

In the next section we show how to determine a unique orthonormal matrix R, i.e., $R^tR = I_3$, given $(r_{13}, r_{23}, r_{31}, r_{32})$ and r_{33} obeying: $0 < r_{13}^2 + r_{23}^2 = r_{31}^2 + r_{32}^2 = 1 - r_{33}^2 \leq 1$. Thus there exist two candidates for the rotation matrix $R(S)$ since $(r_{13}, r_{23}, r_{31}, r_{32})(s_{12}, s_{23}, s_{31}, s_{32})$ is uniquely determined only up to a sign even though $r_{33}(s_{33})$ is uniquely determined. The right one could be determined by the condition: $\det(R) = 1$ ($\det(S) = 1$). The other one is a reflection matrix and given by

$$\begin{bmatrix} I_2 & 0 \\ 0 & -1 \end{bmatrix} R \left(\begin{bmatrix} I_2 & 0 \\ 0 & -1 \end{bmatrix} S\right).$$

IV. Solving a Unique Orthonormal Matrix R Given r_{13}, r_{23}, r_{31}, r_{32}, r_{33} with $0 < r_{13}^2 + r_{23}^2 = r_{31}^2 + r_{32}^2 = 1 - r_{33}^2 \leq 1$

Using the identity (12) again and noticing

$$(r_{13}, r_{23})\bar{R} = (r_{11}r_{13} + r_{21}r_{23}, r_{12}r_{13} + r_{22}r_{23}, r_{13}^2 + r_{23}^2) = (-r_{33}r_{31}, -r_{33}r_{32}, r_{13}^2 + r_{23}^2) \tag{32}$$

which is because $r_{11}r_{13} + r_{21}r_{23} + r_{31}r_{33} = 0$, $r_{12}r_{13} + r_{22}r_{23} + r_{32}r_{33} = 0$, we obtain

$$\begin{bmatrix} -r_{23} & r_{13} \\ r_{13} & r_{23} \end{bmatrix} \bar{R} = \begin{bmatrix} r_{32} & -r_{31} & 0 \\ -r_{33}r_{31} & -r_{33}r_{32}r_{12}^2 & +r_{23}^2 \end{bmatrix} \tag{33}$$

which leads to

$$\bar{R} = \frac{1}{r_{13}^2 + r_{23}^2} \cdot \begin{bmatrix} -r_{23} & r_{13} \\ r_{13} & r_{23} \end{bmatrix} \begin{bmatrix} r_{32} & -r_{31} & 0 \\ -r_{33}r_{31} & -r_{33}r_{32}r_{13}^2 & +r_{23}^2 \end{bmatrix} \tag{34}$$

$$R = \begin{bmatrix} \bar{R} \\ r_3 \end{bmatrix} \tag{35}$$

where $r_3 = (r_{31}, r_{32}, r_{33})$.

We can also directly verify that the matrix R computed by (34) and (35) is orthonormal, i.e., $R^tR = I_3$.

In the next section we give the algorithm to determine R, S, a_3 under the assumptions (10) and (11).

V. Algorithm to Solve R, S, a_3 Under Assumptions (10) and (11)

Now we are ready to give the following simple algorithm:

Step 1) Solve four unit 1×2 vectors u, v, $\bar{u}$, $\bar{v}$ by (18) and (19).

Step 2) Solve β/α, r_{33}, s_{33} by (27).

Step 3) Determine $(r_{13}, r_{23}, r_{31}, r_{32})$ and $(s_{13}, s_{23}, s_{31}, s_{32})$ by (30) and (31).

Step 4) Determine the rotation matrix R by using (34) and (35) and using the condition $\det(R) = 1$. Similarly determine S.

Step 5) Determine a_3 by

$$a_3 = \frac{1}{r_{13}^2 + r_{23}^2}[(r_{13}, r_{23})\bar{B} + r_{33}(r_{31}, r_{32})\bar{A}]. \tag{36}$$

The fact that the solution method is simple does not mean that it will be robust. Thus a stability analysis of the algorithm is needed. Since the authors believe that this research may have a potential for wide applications, they are now furthering their investigation in order to develop a robust version of the algorithm. They will be happy to furnish real implementation results at that time.

VI. Summary

Given two orthographic views the motion and structure of a rigid object cannot be determined no matter how many point correspondences are used, as opposed to the intuitive thought (see Marr [7]). As a matter of fact, from the argument in Section III, we saw both (r_{13}, r_{23}) and (r_{31}, r_{32}) can only be determined up to an arbitrary factor α, $0 < |\alpha| \leq 1$, from two orthographic views. And for each α, $0 < |\alpha| \leq 1$, (20) and (34)–(36) determine two orthonormal matrices when letting $r_{33} = \pm\sqrt{1-\alpha^2}$. One of them is a rotation matrix. Thus the number of solutions is uncountably infinite as α varies. As reported in Aloimonos and Papageorgiou [8], however, using the regularization technique can help to obtain a unique solution in the case of two orthographic views.

Given three orthographic views, the motion and structure can be uniquely determined when assuming (10) and (11) and

$$\text{Rank}\,[(u\bar{B})^t, (v\bar{A})^t, (\bar{v}\bar{A})^t] = 3. \tag{37}$$

As easily seen, assumptions (10) and (11) imply assumption (8). However, it is not very clear whether assumption (37) is always assured by assumptions (10) and (11). If not, then assumptions (10), (11), and (37) comprise a set of necessary and sufficient conditions to uniquely determine the motion and structure from three orthographic views, as clear from the argument in Section III. By comparison, Ullman's method is nonlinear. As commonly happens with a nonlinear technique, Ullman's analysis does not include the uniqueness and convergence verifications.

Acknowledgment

The authors are very thankful to the reviewers for comments and suggestions.

References

[1] S. Ullman, *The Interpretations of Visual Motion.* Cambridge, MA: MIT Press, 1979.

[2] J. Aloimonos and C. M. Brown, "Perception of structure from motion," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition* (Miami Beach, FL, June 22–26, 1986), pp. 510–517.

[3] C. H. Lee, "Structure from motion: An augmented problem and a new algorithm," submitted for publication.

[4] C. Longuet-Higgins, "A computer algorithm for reconstructing a scene from two projections," *Nature,* vol. 293, pp. 133–135, Sept. 1981.

[5] R. Y. Tsai and T. S. Huang, "Uniqueness and estimation of 3-D motion parameters of rigid objects with curved surfaces," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. PAMI-6, no. 1, pp. 13–27, Jan. 1984.

[6] X. Zhuang, T. S. Huang, and R. M. Haralick, "Two-view motion analysis: A unified algorithm," *J. Opt. Soc. Amer., A,* vol. 3, no. 9, pp. 1492–1500, Sept. 1986.

[7] D. Marr, *Vision.* San Francisco, CA: W. H. Freeman and Co., 1982.

[8] J. Aloimonos and A. Papageoriou, "On the kinetic depth effect: Lower bounds, regularization and learning," Center for Automation Research, Univ. of Maryland, College Park, MD 20742, TR CAR-TR-261, Jan. 1987.