論文 89 —26—11—28

Linear Algorithm for Motion Estimation from Point Correspondences over Two Perspective Views : How to Handle Degenerate Cases

(선형해석법에 의한 점대응 움직임 추정 연구 : Degenerate Case에 대한 대책에 관하여)

沈 英 錫*
(Young Serk Shim)

要 約

두 프레임간에 걸친 점대응으로부터 삼차원 강체의 움직임 및 구조를 추정하는 문제는 기본적으로 비선형이나, 대응수가 증가하면 선형해석을 통하여 풀 수가 있다. 그러나 대응점들이 삼차원 공간상에서 특정한 기하학적 조건을 만족하거나, 그에 가까운 분포를 이루는 경우에는 선형해석법에 의한 움직임 추정이 불가능해진다. 본 논문에서는 이와같이 degenerate한 경우에도 저차다항식에 의해 표현되는 보조식의 해를 구함으로써 움직임 추정이 가능함을 보인다.

Abstract

For determining motion/structure of a 3-D rigid object from point correspondences over two perspective views, a linear algorithm was developed in Refs. 3 and 4. This algorithm fails when the 3-D points under observation satisfy certain geometrical constraints, as demonstrated in Refs. 7 and 8. In the present paper, we show that the linear algorithm can be resurrected in these degenerate cases by adding additional lower order polynomial constraints to original linear equations.

I. Introduction

The problem of determining 3-D motion and structure from image sequences is an important task in computer vision, and has attracted much attention. It has a broad range of practical applications such as autonomous vehicle navigation, robot guidance, and monitoring industrial process.

*正會員, 慶北大學校 電子工學科
(Dept. of Elec. Eng., Kyungpook Nat'l Univ.)
接受日字 : 1989年 4月 10日

To solve such a problem, two basic approaches are frequently used: methods based on an optical flow, and methods based on line or point correspondences. In the latter feature based approach, each frame of image sequence is segmented first, and the feature points/lines resulting from the projection of vertices and edge of the object surfaces are marked. Next the correspondence of these features between successive frames is established. As a final step, the 3-D motion parameters and object structure are derived. For the flow based approach, see[1,2].

Much work has been devoted to the final motion estimation step, to derive the motion parameters and object structure from point correspondences over two perspective views. In particular, a linear algorithm has been developed by Longuet-Higgins[3], and Tsai and Huang[4]. This algorithm requires 8 (or more) point correspondences while basic nonlinear approaches[5,6] using iterative techniques need 5 (or more) correspondences. A necessary and sufficient condition, a quadratic surface condition, for this algorithm to fail has also been established[7,8]. In this paper, we show that in these degenerate cases where the above condition holds, the linear algorithm can nevertheless be resurrected. The modified procedure requires, in addition to solving the 8 (or more) linear equations, the least squares solution of three polynomial equations.

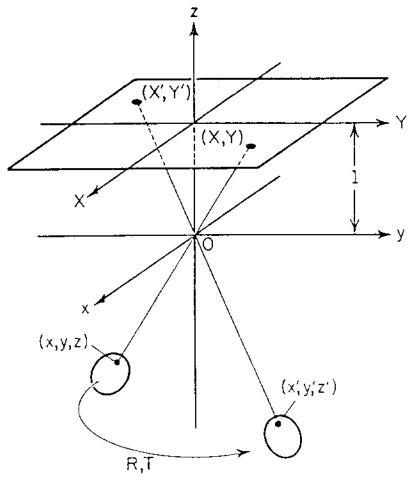

Fig.1. Basic imaging geometry.

II. Review of the 8-Point Algorithm

1. Problem Statement

We shall use the following notations (see Fig. 1). The object-space coordinates are denoted by lowercase letters, and the image-plane coordinates by uppercase letters. The coordinates of points before the motion are unprimed, those after the motion are primed. Thus

(x,y,z) = coordinates of object point before motion

(x', y', z') = coordinates of object point before motion

(X,Y) = coordinates of image point after motion

(X', Y') = coordinates of image point after motion

We assume that the image plane is at z = 1, the origin of the image plane is on the z-axis, and the X- and Y-axis are parallel to the x- and y- axes,- respectively. The point (X,Y) is the central projection of the point (x,y,z) with respect to the origin (0,0,0).

From kinematics,

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = R \begin{bmatrix} x \\ y \\ z \end{bmatrix} + T \qquad (1)$$

where R is a rotation (right-handed orthonormal) matrix and T a translation vector

$$T = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} \qquad (2)$$

Our problem is:
Given-N image point correspondences

$$(X_i,\ Y_i) \longleftrightarrow (X_i',\ Y_i'),\quad i=1,2,\cdots,N.$$

Find- The motion parameters T (to within a scale factor) and R.

2. The 8-Point Linear Algorithm

The following projective equations relate the 3-D spatial coordinates and their coresponding 2-D image coordinates:

$$X = x/z, \qquad Y = y/z,$$
$$X' = x'/z', \quad Y' = y'/z' \qquad (3)$$

Recall that the 3-D coordinates of a point before and after the motion are related by(1). Taking any nonzero vector which is colinear with T and taking its cross-product with both sides of (1), we obtain

$$\frac{z'}{z} T \times [X' \ Y' \ 1]^t = T \times R [X \ Y \ 1]^t \qquad (4)$$

and after taking dot product of both sides of (4) with [X' Y' 1],

$$[X' \ \ Y' \ \ 1] \quad [T \times R] \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = 0 \qquad (5)$$

where $T \times R \triangleq [T \times r_1 \ \ T \times r_2 \ \ T \times r_3]$; r_1, r_2, r_3 being the column vectors of R^t.

Now if we define

$$E = \begin{bmatrix} e_1 & e_2 & e_3 \\ e_4 & e_5 & e_6 \\ e_7 & e_8 & e_9 \end{bmatrix} \triangleq GR \qquad (6)$$

where

$$G = \begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{bmatrix} \qquad (7)$$

then it can be readily shown that

$$T \times R = GR \qquad (8)$$

and

$$[X' \ \ Y' \ \ 1] \ E \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = 0 \qquad (9)$$

which is linear and homogeneous in the nine unknowns $e_1, e_2, \ldots, e_9$.

From the N point correspondences, we have

$$B \begin{bmatrix} e_1 \\ e_2 \\ \cdot \\ \cdot \\ \cdot \\ e_9 \end{bmatrix} = 0 \qquad (10)$$

where

$$B \triangleq \begin{bmatrix} X_1'X_1 & X_1'Y_1 & X_1' & Y_1'Y_1 & Y_1' & X_1 & Y_1 & Y_1 & 1 \\ \cdot & & & & & & & & \cdot \\ \cdot & & & & & & & & \cdot \\ \cdot & & & & & & & & \cdot \\ X_N'X_N & X_N'Y_N & X_N' & Y_N'X_N & Y_N'X_N & Y_N' & X_N & Y_N & 1 \end{bmatrix} \qquad (11)$$

Assuming the rank of B is 8, we can solve (10) to get k E, where k is an unknown scale factor.

Once k E is obtained, several methods are available [3,4,9] to decompose it uniquely to find T (to within a scale factor) and R.

3. Degenerate Cases

The linear algorithm fails if the rank of B is less than 8 (in which case we say that degeneracy occurs). A necessary and sufficient condition for degeneracy can be stated geometrically if we assume that the object is stationary and the camera is moving[6]. Let the origin of the camera system be 0 and 0′ before and after the motion, respectively. Then, assuming $T \neq 0$, the rank of B is less than 8 if and only if the N 3-D points lie on a quadratic surface passing through 0 and 0′.

III. How to Resurrect the 8-point Linear Algorithm in Degenerate cases.

1. Basic Idea

It is important to note that although the linear algorithm as described in Section II.B fails if Rank (B) < 8, it does not necessarily mean that a point or unique solution to the motion parameters cannot be found by other methods. In fact, we shall show presently that the linear algorithm with appropriate modifications can be used even in the degenerate cases.

The basic idea is as follows. Let

$$\text{Rank}(B) = M < 8 \qquad (12)$$

Then there are

$$K = 9 - M \qquad (13)$$

linearly independent vectors in the null space of B. A set of these vectors can be found by solving (10). Putting each of these vectors in the form of a 3 x 3 matrix, we obtain a set of K E-matrices which we shall call E_i;i=1,2,..., K: Then a general solution to (10) is

$$E = \sum_{i=1}^{K} a_i E_i \qquad (14)$$

where the a_i's are arbitrary constants.

The thing left to do is to find values of a_i's such that E is in the form of (6), i.e., it is decomposable into the product of a skew matrix and a right-handed orthonormal matrix.

2. Decomposability Conditions

We show that the decomposability conditions for E can be expressed in the form of three simultaneous polynomial equations with the a_i's as unknowns.

Let

r_i = i-th row of R

ϵ_i = i-th row of E (i=1,2,3)

Then (6) can be written as

$$\epsilon_1 = -t_3 r_2 + t_2 r_3 \qquad (15\text{-a})$$

$$\epsilon_2 = t_3 r_1 - t_1 r_3 \qquad (15\text{-b})$$

$$\epsilon_3 = -t_2 r_1 + t_1 r_2 . \qquad (15\text{-c})$$

Our problem is to find conditions on $\epsilon_1, \epsilon_2, \epsilon_3$ such that they can be expressed as in (15), where t_1, t_2, t_3 are real numbers and $\begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}$ is a right-handed orthornormal matrix. These conditions can then be expressed in terms of the a_i's via (14).

Lemma Any two 3-vectors ϵ_1 and ϵ_2 can be expressed as

$$\epsilon_1 = -t_3 r_2 + t_2 r_3 \qquad (16)$$

$$\epsilon_2 = t_3 r_1 - t_1 r_3 \qquad (17)$$

where t_1, t_2, t_3 are real numbers and $\begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}$ is a right-handed orthonormal matrix.

[Proof] From (16) and (17),

$$\epsilon_1 \times \epsilon_2 = t_3 (t_1 r_1 + t_2 r_2 + t_3 r_3) \qquad (18)$$

Eqs. (16), (17), (18) can be written as

$$A \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_1 \times \epsilon_2 \end{bmatrix} \qquad (19)$$

where

$$A \triangleq \begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ t_3 t_1 & t_3 t_2 & t_3^2 \end{bmatrix} \qquad (20)$$

Case 1- If $\epsilon_1 \times \epsilon_2 = 0$, then we can choose r_3 as the normalized vector of ϵ_1 or ϵ_2. Then, $t_3=0$. and we do not have to worry about r_1 and r_2.

Case 2- If $\epsilon_1 \times \epsilon_2 \neq 0$, then

$$\det(A) = \| \epsilon_1 \times \epsilon_2 \|^2 \neq 0.$$

We solve (19) to get

$$r_3 = [t_3^2 (\epsilon_1 \times \epsilon_2) + t_2 t_3^2 \epsilon_1 - t_1 t_3^2 \epsilon_2] / \det(A)$$

$$r_2 = [t_2 t_3 (\epsilon_1 \times \epsilon_2) - t_3 (t_1^2 + t_3^2) \epsilon_1 - t_1 t_2 t_3 \epsilon_2] / \det(A)$$

$$r_1 = [t_1 t_3 (\epsilon_1 \times \epsilon_2) + t_1 t_2 t_3 \epsilon_1 + t_3 (t_2^2 + t_3^2) \epsilon_2] / \det(A)$$

$$(21)$$

It can be readily verified that r_1, r_2, r_3 form a right-handed orthonormal basis, if t_1, t_2, t_3 satisfy:

$$t_2^2 + t_3^2 = \| \epsilon_1 \|^2$$

$$t_1^2 + t_3^2 = \| \epsilon_2 \|^2$$

$$t_1 t_2 = -(\epsilon_1 \cdot \epsilon_2) \qquad (22)$$

[End of Proof]

Now we proceed to find conditions on ϵ_1, ϵ_2' ϵ_3 such that ϵ_3 can be expressed as in (15C).

Let

$$\epsilon_3 = \mu r_1 + \nu r_2 + \omega r_3 \qquad (23)$$

Then we want to find conditions such that (23) has

$$(\mu, \nu, \omega) = (-t_2, t_1, 0) \qquad (24)$$

as a solution. We derive this decomposability conditions by expressing the magnitude and direction of ϵ_3 in terms of $\epsilon_1, \epsilon_2, \epsilon_3$.

From (15), the direction of ϵ_3 can be determined by:

$$\epsilon_3 \cdot (\epsilon_1 \times \epsilon_2) = 0 \qquad (25)$$

and

$$2(\epsilon_3 \cdot \epsilon_1)^2 + 2(\epsilon_3 \cdot \epsilon_2)^2 = \| \epsilon_3 \|^2 (\| \epsilon_1 \|^2 + \| \epsilon_2 \|^2 - \| \epsilon_3 \|^2) \qquad (26)$$

From (15C), the magnitude of ϵ_3 is determined by

$$\| \epsilon_3 \|^2 = t_2^2 + t_1^2 \qquad (27)$$

which becomes, with the help of (22):

$$\| \epsilon_3 \|^4 = (\| \epsilon_2 \|^2 - \| \epsilon_1 \|^2)^2 + 4(\epsilon_1 \cdot \epsilon_2)^2 \qquad (28)$$

Eqs. (25), (26), and (28) are the desired decomposability conditions. To check, we rewrite them in terms of $\mu, \nu, \omega, t_1, t_2, t_3$ using (15A), (15B), and (23):

$$t_3(t_1\mu + t_2\nu + t_3\omega) = 0 \tag{29}$$

$$(t_1^2 + t_2^2 + t_3^2)\omega^2 = 0 \tag{30}$$

$$\mu^2 + \nu^2 + \omega^2 = t_1^2 + t_2^2 \tag{31}$$

Solving, we get the desired solution

$$(\mu, \nu, \omega) = (-t_2, t_1, 0) \tag{32}$$

or a false solution

$$(\mu, \nu, \omega) = (t_2, -t_1, 0) \tag{33}$$

corresponding to a left-handed orthonormal coordinates, and can be automatically discarded in the decomposition process. Therefore we can say that our decomposability conditions is neceesary and sufficient for the motion to be described as a translation after a rotation in a right-handed orthonormal coordinates.

Note that via (14), Eqs. (25), (26), (28) can be written in terms of the a_i's. These are polynomial equations of degrees 3,4, and 4, respectively. The equations are homogeneous, reflecting the fact that the a_i's can only be determind to within an unknown scale factor. We can conceptually setting one of the a_i's to 1, and thus: if the number of a_i's is K, the number of unknowns in the three equations is (K-1). The interesting cases are:

The case of M=8 is the nondegenerate case where the original linear alogirhtm works and the solution to motion/structure is unique. When M=7 or 6, we have more equations than unknowns, so the solution to the a_i's is very likely unique. When M=5, the number of equations is equal to the number of unknowns. Since the equations are nonlinear (polynomial), there are likely more than one (but a finite number of) solutions to the a_i's. For each solution to the a_i's, we obtain an E (using (10)) to within a scale factor, from which we get a solution to the motion and structure. For M $\leqslant$ 4, the number of equations becomes smaller than the number of unknowns. Therefore, the number of solutions to the a_i's (and hence to motion and structure) becomes infinite. In fact, the solution space is continuous.

For the cases M=7,6 and 5, a good way to solve the three polynomial equations for the a_i's is to find a least squares solution subject to the constraint.

$$\sum_{i=1}^{K} a_i^2 = 1 \tag{34}$$

Iterative methods can be used to obtain the solution. Since we do not know a priori the number of solutions, global search is needed. The computation can be tedious, especially for M=5.

3. Summary of the Algorithm

Step 1: By solving (6), find a set of basis vectors of the null space of the matrix B. If Rank(B)=M, then the number of such vectors is K=9-M.

Step 2: Using (10), express ϵ (i=1, 2, ..., 9) in terms of a_i (i=1,2, ..., K).

Step 3: Construct the three polynomial equations (21), (22), and (24); and solve them for the a_i's.

Step 4: For each solution of the a_i's, get E from (10) and decompose it (using the technique of Ref. 3,4, or 9) to obtain the motion parameters and object structure.

IV. Computer Simulation Results.

1. Seven Points in General Positions, Rank(B) =7

In every case we have tried, the solution is unique. We shall give one example here.

Ground truth: 3-D object coordinates of 7 points before motion-

(2,2,2), (3,1,3), (-2,2,2), (2,-2,3)
(-1,-3,3.5), (-4,-3,2.5), (3,0,3)

Rotation-
Direction of rotation axis = (1,1,1)
Angle of rotation = 30°
Translation = (1,0,1)
Given point correspondences:

i	X_i	Y_i	X'_i	Y'_i
1	1	1	1	0.666666
2	1	0.3333333	1.34641	0.3535899
3	−1	1	−0.1616507	0.1676698
4	0.6666667	−0.6666667	1.672028	−0.7320508
5	−0.2857143	−0.8571429	5.93637	−1.142226
6	−1.6	−1.2	−0.3312093	−1.437365
7	1	0	1.577335	0.089316411

Solution:

Step 1- Rank(B) =7

$$E_1 = \begin{bmatrix} 0.22912069 & 1 & 0.1671754 \\ -1.101394 & 0.5376813 & 0.487313 \\ -0.5702385 & -0.7751201 & 0 \end{bmatrix}$$

$$E_2 = \begin{bmatrix} 0.510991 & 0 & -1.623915 \\ -0.6215896 & 0.3593681 & 0.5473321 \\ 0.7621039 & -0.8392533 & 1 \end{bmatrix}$$

Step 2- $E=a_1E_1 + a_2E_2$ Thus, $e_1=0.2291069a_1 + 0.510991\, a_2$, etc.

Step 3- Since in this simulation the data are essentially noise-free, instead of finding the least-squares solution of Eqs. (21), (22), (24), we take an easier route. We first construct Eq. (21) and set $a_2 = 1$:

$$0.02635941a_1^3 - 0.01565024a_1^2 + 0.01741159a_1 + $$
$$+ 0.01595911 = 0$$

The three roots of this equation are:

$-0.582126,\ 3.734082,\ 2.785306$

For each root we get E via (10) and substitute it into Eqs. (22) and (24). Only the second root satisfies these two equations. Thus, we choose

$a_1 = 3.734082$

Step 4-

$$E = a_1E_1 + E_2 = \begin{bmatrix} 1.366495 & 3.734082 & -0.9996085 \\ -4.734287 & 2.367114 & 2.36699999 \\ -1.367176 & -3.733616 & 1 \end{bmatrix}$$

Finally, using the method of Ref. 9 we decompose E to get:

Rotation axis direction
= (0.5772734, 0.5773939, 0.5773835)

Rotation angle = 29.99962°

Translation = (4.099894, -0.000624, 4.100011)

Note that the translation is correct only to within a scale factor as was already mentioned in the problem statement. To determine the scale parameter, range information of at least one point should be provided in separate way.

2. Eight Point on Vertices of A Cube.

This case is degenerate as stated in Ref. 1. We found that Rank(B) = 7 and that the solution to motion/structure is unique.

3. Six Points in General Positions, Rank(B)=6.

In every case we have tried, the solution to motion/structure is unique.

4. Five Points in General Positions. Rank(B)=5

In each case we have tried, the number of solution to motion/structure is 2.

V. Conclusion

In this paper, we have presented an algorithm to resurrect the well known linear algorithm for the two view motion analysis based on the point correspondences even in degenerate cases. Our method comes from the derivation of the decomposability condition which is neceessary and sufficient for the motion to be described by a translation after a rotation in a right-handed orthonormal coordinates. Computer simulation results have confirmed the theoretical derivation. The decomposability condition expressed by three 3-or 4-th order polynomial equations could be used for the seek of least squares solution when the image points are measured inaccurately and/or overdetermined.

References

[1] B.G. Schunck, "Image flow: Fundamentals and future research," IEEE International Conference on Computer Vision and Pattern Recognition, pp. 560-571, 1985.

[2] X. Zhuang, R.M. Haralick, and Y. Zhao, "From depth and optical flow to rigid body motion," IEEE International Conference on Computer Vision and Pattern Recognition, pp. 393-397, 1988.

[3] H.C. Longuet-Higgins, "A computer algorithm for reconstructing a scene from two projections," Nature, vol. 293, pp. 133-135, September 1981.

[4] R.Y. Tsai and T.S. Huang, "Uniqueness and estimation of 3-D motion parameters of rigid bodies with curved surfaces," *IEEE Transactions PAMI,* vol. PAMI-6, no. 1, pp. 13-27, January 1984.

[5] J.W. Roach and J.K. Aggarwal, "Determining the movement of objects from a sequence of images," *IEEE Transactions*

PAMI, vol. PAMI-2, no. 6, pp. 554-562, November 1980.

[6] J.Q. Fang and T.S. Huang, "Some experiments on estimating the 3-D motion parameters of a rigid body from two consecutive image frames," *IEEE Transactions PAMI,* vol. PAMI-6, no. 5, pp. 545-554, September

[7] H.C. Longuet-Higgins, "The reconstruction of scene from two projections: configurations that defeat the 8-point algorithm," Proc. The First Conference on Artificial Intelligence Applications, pp. 395-397, December 1984.

[8] X. Zhuang, T.S. Huang, and R.M. Haralick, "Two-view motion analysis: A unified algorithm," Journal of Optical Society of America A, Optics an Image Science, vol. 3, no. 9, pp. 1492-1500, September 1986.

[9] B.L. Yen and T.S. Huang, "Determining 3-D motion parameters of a rigid body: A vector-geometric approach," Proc. the ACM Workshop on Motion, April 1983.

著者紹介

沈 英 錫 (正會員) 第26卷 第2號 參照
현재 경북대학교 전자공학과
부교수