TWO VIEW MOTION ANALYSIS UNDER A SMALL PERTURBATION

Xinhua Zhuang*, Robert M. Haralick
Machine Vision International
Ann Arbor, MI 48104

*Visiting Scientist from People's Republic of China

Abstract

Given a set of corresponding points from a moving object is two perspective projection images.

The paper completely solves the two view motion problem. We show how to use the corresponding point set to determine mode of motion, rotation, translation orientation and relative depths. Also we give a noise robust algorithm which works well under small perturbations.

I. Introduction

It is well known that the two view motion analysis has a special and basic importance in robotic vision. Lots of contributions have been made in recent years ([1],[2],[3]). Unfortunately, most of researches neglect, among other things, importance of exploring the general solution of Two View - Motion Equation and differentiating modes of motion. It leads to imperfect or even incorrect results. Without a satisfactory theory it is impossible to develop a sound algorithm.

This paper briefly reviews results obtained recently by authors [1] which comprise a complete solution to the two view motion problem. Besides, a noise robust algorithm which works well under small perturbations is included.

II. General Solution of Basic Two View - Motion Equation. Surface Assumption

Assume that a rigid body is in motion in the half-space $z<0$. Take a particular point p on the object. Let (x,y,z) be the spatial coordinates of p before the motion and (x',y',x') be the coordinates after the motin. Let (X,Y) be central projective coordinates of p before the motion (and (X',Y') after the motion) onto the plane z=1 with the projective center at origin 0. The following projection equations relate the 3-D point coordinates and corresponding 2-D projective point coordinates:

$$X = x/z,\ Y = y/z$$

$$X' = x'/z',\ Y' = y'/z'$$

As known, any 3-D rigid body motion M is equivalent to a rotation R_o followed by a translation T_o such that

$$(x',y',z')^t = R_o\,(x,y,z)^t + T_o$$

where R_o is a 3X3 orthonormal matrix with det $(R_o)=1$ and T_o is a 3X1 vector. 't' represents the transposition operation. From the motion equation it is easy to obtain

$$T_o \times (X',Y',1)^t = T_o \times [R_o(X,Y,1)^t]$$

and hence the following Two View - Motion Equation

$$(X',Y',1)\ \{T_o \times [R_o(X,Y,1)^t]\} = 0 \qquad (1)$$

Let

$$T_o = (\Delta x_o, \Delta y_o, \Delta z_o)^t$$

$$G_o = \begin{bmatrix} 0 & -\Delta z_o & \Delta y_o \\ \Delta z_o & 0 & -\Delta x_o \\ -\Delta y_o & \Delta x_o & 0 \end{bmatrix}$$

It is easy to verify that for any 3X1 vector v there holds

$$T_o \times v = G_o v$$

As a result, Two View - Motion Equation turns out to be a familiar form

$$(X',Y',1)\ G_o\ R_o\ (X,Y,1)^t = 0$$

which is established in [2], [3]. Let

$$E_o = G_o R_o$$

Now it is clear that for any projection point correspondence pair [(X,Y,), (X',Y')] the following equality always holds:

$$(X',Y',1)\ E_o\ (X,Y,1)^t = 0$$

Conversely, if a real 3X3 matrix E satisfies

$$(X',Y',1)\ E\ (X,Y,1)^t = 0 \qquad (2)$$

for a set of projection point correspondence pairs, denoted by P, then we have (see [1])

<u>Theorem 1.</u> $E = \alpha E_o$ (α any real number) when $T_o \neq 0$ or GR_o (G any skew symmetric matrix) when $T_o =$

0 if and only if Surface Assumption holds, that is, the surface patch or the group of surface points S which produces P cannot be contained in a quadratic surface of form

$$(x,y,z) \, U \, (x,y,z)^t + v^t(x,y,z)^t = 0$$

with $||U+U^t|| + ||v|| = 0$ and $T^t R_o U = v^t$.

Let

$$A = (XX', YX', X', XY', YY', Y', X, Y, 1)$$

$$W = \sum A^t A \ (\geqslant 0, \text{ as easily seen})$$

$$h = (h_1, h_2, \ldots, h_9)^t$$

$$E = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix}$$

Then, it is proved (see [1]) that E satisfies with (2) for $[(X,Y),(X',Y')] \in P$ if and only if h satisfies with

$$Wh = 0 \quad (3)$$

Theorem 2. (see [1]) Under Surface Assumption there are 8 element largest linear independent A's denoted by $A_1,\ldots,A_8$ such that

$$\text{Rank } (W) = \text{Rank } (\sum A_i^t A_i) = 8$$

and hence the general solution h is one parameter if and only if $T_o \neq 0$ and there are 6 element largest linear independent A's denoted by $A_1,\ldots,A_6$ such that

$\text{Rank } (W) = \text{Rank } (\sum A_i^t A_i) = 6$ and hence the general solution h is three parameters if and only if $T_o = 0$.

As a result of Theorem 2 at least 6 (8) point pairs must be contained in P depending on whether or not T_o is zero. More projection point correspondence pairs are preferable to guarantee Surface Assumption and to smooth out any noise effects. The general solution h (or the same, E) of (3) has two and only two decompositions (see [1],[2]):

$$E = TxR = (-T)xR' \quad (4)$$

where both R and R' are orthonormal matrices of the first kind. One of them equals R_o, and depending on whether or not T_o is zero T is any real vector or equals αT_o with α any real number. (Note: TxR $=[Txr_1, Txr_2, Txr_3]$). A direct procedure for decomposing E is given in [1] and Section IV.

III. Determination of Mode of Motion, Rotation, Translation Orientation and Relative Depth

Theorem 3 (Mode of Motion) (see[1]) Assume that $M=(T_o,R_o)$ and $E(=TxR=(-T)xR')$ is a nonzero solution of Two View - Motion Equlation. Then, $(T_o,R_o) = (o,R)$ holds if and only if for any two projection point correspondence pairs $[(X_i,Y_i), (X'_i,Y'_i)]$ (i=1,2) there hold

$$v'_i / ||v'_i|| - Rv_i / ||v_i|| = o \quad (i=1,2) \quad (5)$$

where

$$v'_i = (X'_i, Y'_i, 1)^t \text{ and } v_i = (X_i, Y_i, 1)^t$$

Theorem 4 (Rotation and Translation Orientation when $T_o \neq o$) (see [1])

Assume that $M = (T_o,R_o)$ with $T_o \neq 0$ and

E (=T x R = (-T) x R') is a non zero solution of Two View - Motion Equation. Then, $R_o = R$ and $T_o / ||T|| = \pm T / ||T||$ hold if and only if for any two projection joint correspondence pairs $[(X_i,Y_i),(X'_i,Y'_i)]$ (i=1,2) there hold

$$||T \times Rv_i|| \, v'_i - || T \times v'_i || \, Rv_i \pm ||v'_i \times Rv_i|| \, T = o$$

$$(i=1,2) \quad (6)$$

Theorem 5. (Relative Depth) (see[1]) Assume that $M = (T_o,R_o)$ and $E(=TxR=(-T)xR')$ is a non zero solution of Two View-Motion Equation. Then the relative depth is given by:

$$z'/z = ||v|| / ||v'|| \quad \text{when } T_o = 0 \quad (7)$$

$$z'/z = ||T \times R_o v|| / ||T \times v'|| \quad \text{when } T_o \neq 0$$

Let

$$\begin{cases} A(v,v',R) = ||v'/||v'|| - Rv/ \, ||v|| \, || \\ H(v,v',T,R) = || \, ||TxRv|| \, v' - ||Txv'||Rv + ||v'xRv||T|| \end{cases} \quad (8)$$

A more efficient and even stronger procedure to detemrine the mode of motion, the rotation and the translation orientation comes out:

Theorem 6. (See [1]) Assume that $M =(T_o,R_o)$ and E $(= TxR = (-T) \times R')$ is a nonzero solution of Two View-Motion Equation. Then

$(T_o,R_o) = (0,R)$ holds if and only if for any $n \ (\geqslant 2)$ pairs $[(X_i,Y_i),(X'_i,Y'_i)] \in P$ (i = 1,2,...n)

$$\sum_{i=1}^{n} A(v_i,v'_i,R) < \sum_{i=1}^{n} A(v_i,v'_i,R'),$$

$$\sum_{i=1}^{n} H(v_i,v'_i,T,R), \quad \sum_{i=1}^{n} H(v_i,v'_i,-T,R), \quad (9)$$

$$\sum_{i=1}^{n} H(v_i,v'_i,-T,R'), \quad \sum_{i=1}^{n} H(v_i,v'_i,T,R')$$

and

$(T_o/||T_o||,R_o) = (\pm T/||T||,R)$ holds if and only if

$$\sum_{i=1}^{n} H(v_i,v'_i, \pm T,R) < \sum_{i=1}^{n} A(v_i,v'_i,R),$$

$$\sum_{i=1}^{n} A(v_i,v'_i,R'), \quad \sum_{i=1}^{n} H(v_i,v'_i \mp T,R) \quad (10)$$

$$\sum_{i=1}^{n} H(v_i,v'_i,-T,R'), \quad \sum_{i=1}^{n} H(v_i,v'_i,T,R')$$

Theorem 6 is important since, based on it, a noise robust algorithm is developed. See next section.

IV. Algorithm

Step 1. Solve min $h^t W h$ with $||h|| = 1$

Step 2. Let

$$L_1 = [h_1, h_2, h_3]$$

$$L_2 = [h_4, h_5, h_6]$$

$$L_3 = [h_7, h_8, h_9]$$

$$E = \begin{bmatrix} L_1 \\ L_2 \\ L_3 \end{bmatrix}$$

Step 3. Let

$$\alpha = (||L_2||^2 + ||L_3||^2 - ||L_1||^2)/2$$

$$\beta = (||L_3||^2 + ||L_1||^2 - ||L_2||^2)/2$$

$$\gamma = (||L_1||^2 + ||L_2||^2 - ||L_3||^2)/2$$

Step 4. If $|\alpha| \geqslant |\beta|, |\gamma|$, then let

$$T = \begin{bmatrix} \sqrt{\alpha} \\ -\langle L_1, L_2 \rangle / \sqrt{\alpha} \\ -\langle L_1, L_3 \rangle / \sqrt{\alpha} \end{bmatrix}$$

$$r_1 = [(L_2 x L_3) x L_1 + \sqrt{\alpha}(L_2 x L_3)]/(||T||^2 \sqrt{\alpha})$$

$$r'_1 = [(L_2 x L_3) x L_1 - \sqrt{\alpha}(L_2 x L_3)]/(-||T||^2 \sqrt{\alpha})$$

$$r_2 = (r_1 x L_2)/\sqrt{\alpha}$$

$$r'_2 = -(r_1 x L_2)/\sqrt{\alpha}$$

$$r_3 = (r_1 x L_3)/\sqrt{\alpha}$$

$$r'_3 = -(r_1 x L_3)/\sqrt{\alpha}$$

GO TO STEP 7

(Where $\langle .,. \rangle$ represents the scalar product between two row vectors. For convenience the cross product operation 'x' also acts on two row vectors and produces a row vector.)

Step 5. If $|\beta| \geqslant |\gamma|$, then let

$$T = \begin{bmatrix} -\langle L_2, L_1 \rangle / \sqrt{\beta} \\ \sqrt{\beta} \\ -\langle L_2, L_3 \rangle / \sqrt{\beta} \end{bmatrix}$$

$$r_2 = [(L_3 x L_1) x L_2 + \sqrt{\beta}(L_3 x L_1)]/(||T||^2 \sqrt{\beta})$$

$$r'_2 = [(L_3 x L_1) x L_2 - \sqrt{\beta}(L_3 x L_1)]/(-||T||^2 \sqrt{\beta})$$

$$r_3 = (r_2 x L_3)/\sqrt{\beta}$$

$$r'_3 = -(r_2 x L_3)/\sqrt{\beta}$$

$$r_1 = (r_2 x L_1)/\sqrt{\beta}$$

$$r'_1 = -(r_2 x L_1)/\sqrt{\beta}$$

GO TO STEP 7

Step 6. Let

$$T = \begin{bmatrix} -\langle L_3, L_1 \rangle / \sqrt{\gamma} \\ -\langle L_3, L_2 \rangle / \sqrt{\gamma} \\ \sqrt{\gamma} \end{bmatrix}$$

$$r_3 = [(L_1 x L_2) x L_3 + \sqrt{\gamma}(L_1 x L_2)]/(||T||^2 \sqrt{\gamma})$$

$$r'_3 = [(L_1 x L_2) x L_3 - \sqrt{\gamma}(L_1 x L_2)]/(-||T||^2 \sqrt{\gamma})$$

$$r_1 = (r_3 x L_1)/\sqrt{\gamma}$$

$$r'_1 = -(r_3 x L_1)/\sqrt{\gamma}$$

$$r_2 = (r_3 x L_2)/\sqrt{\gamma}$$

$$r'_2 = -(r_3 x L_2)/\sqrt{\gamma}$$

Step 7. Let

$$R = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}$$

$$R' = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}$$

Step 8. If ($\sum A(v_i, v'_i, R) < \sum A(v_i, v'_i, R')$, $\sum H(v_i, v'_i, T, R)$, $\sum H(v_i, v'_i, -T, R)$, $\sum H(v_i, v'_i, -T, R')$, $\sum H(v_i, v'_i, T, R')$)

then with { the rigid body motion M is a pure rotation

$R_o = R$ and $z'/z = ||v|| / ||v'||$;

GOTO Step 12 }

Step 9. If ($\sum A(v_i, v'_i, R') < \sum A(v_i, v'_i, R)$, $\sum H(v_i, v'_i, T, R)$, $\sum H(v_i, v'_i, -T, R)$, $\sum H(v_i, v'_i, -T, R')$, $\sum H(v_i, v'_i, T, R')$)

then with { the rigid body motion M is a pure rotation

$R_o = R'$ and $z'/z = ||v|| / ||v'||$;

GOTO Step 12 }

Step 10. If ($\sum H(v_i, v'_i, \pm T, R) < \sum A(v_i, v'_i, R)$,

$\sum A(v_i, v'_i, R')$, $\sum H(v_i, v'_i, \mp T, R)$,

$\sum H(v_i, v'_i, -T, R')$, $\sum H(v_i, v'_i, T, R')$)

then{ $(T_0 / ||T_0||, R_0) = (\pm T / ||T||, R)$ and

$z'/z = ||TxRv|| / ||Txv'||$;

GOTO Step 12 }

Step 11. If ($\sum H(v_i, v'_i, \pm T, R') < \sum A(v_i, v'_i, R)$,

$\sum A(v_i, v'_i, R')$, $\sum H(v_i, v'_i, T, R)$,

$\sum H(v_i, v'_i, -T, R)$, $\sum H(v_i, v'_i, \mp T, R')$)

then { $(T_0 / ||T_0||, R_0) = (\pm T / ||T||, R')$ and

$z'/z = ||TxR'v|| / ||Txv'||$ }

Step 12. STOP

V. Conclusion

It seems that two view-motion problem for a single rigid body is completely solved. In [4], [5] optic flow-motion problem for a signle rigid body is completely solved, too. In [6] authors prove equivalence within three problems: two view motion analysis, stereo vision and a moving camera's positioning. Authors expect that not very long two view-multi rigid body motion problem could be solved.

Reference

[1] Xinhua Zhuang and Robert M. Haralick, Two View Motion Analysis (in preparation).

[2] Tsai and Huang, Uniqueness and Estimation of 3-D Motion Parameters of Rigid Objects with Curved Surfaces, IEEE PAMI-6, No. 1 (1984).

[3] Longuet-Higgins, A Computer Algorithm for Reconstructuring a Scene from Two Projections, Nature 293 (Sep. 1981) 133-135.

[4] Xinhua Zhuang and Robert M. Haralick, Rigid Body Motion and Optic Flow Image, The First Conference on Artificial Intelligence Applications. Denver, December 5-7, 1984.

[5] Xinhua Zhuang and Robert M. Haralick, Rigid Body Motion and Optic Flow, (submitted for publication).

[6] Xinhua Zhuang and Robert M. Haralick, Two View Motion Analysis, Stereo Vision and A Moving Camera's Positioning - Their Equivalence and a New Solution Procedure, (submitted for publication).