A Study on Real-Time Photogrammetry for Three-Dimensional Token Tracking

Yongsheng ZHANG

Institute of Remote Sensing & Photogrammetric Engineering

Zhengzhou Institute of Surveying & Mapping

E-mail: yszhang@public2.zz.ha.cn

P. R. China

Commission V, Working Group ICV/Ⅲ

KEYWORDS: real—time photogrammetry, three—dimension, token tracking, stereo vision, motion vision

ABSTRACT

This paper presents a study on real-time photogrammetry for token tracking in three dimensions. We have a sequence of stereo frames from which we will compute, using the algorithms of stereo vision, a sequence of three-dimensional frames. To be concrete, these three-dimensional frames will consist of sets of three-dimensional line segments. By tracking we mean the ability to follow the motion of a given segment and to estimate its kinenatecs. Since much more information is available than in the two-dimensional, we may expect to be able to solve much more difficult problems. Indeed, we will directly estimate the three-dimensional kinematics of the line segments and will be able to cope with the problem of multiple-boject motion. We choose the state and determine the plant and the measurement equations.

1. A BIT OF KINEMATICS

We know from elementery kinematics that the motions of the points of a moving rigid body are conveniently described by a six-dimensional entity called a screw, which is defind at every point p of space and noted $S(P) = (\Omega, V(P))$. Ω is called the angular velocity, and V (P) is the velocity of the point of the solid in motion coinciding with P. The kinematic screw at one point entirely describes the motion of the solid since, at every point M, the velocity of the point of the solid coinciding with M is given by

$$V(M) = V(P) + \Omega \wedge PM \qquad (1)$$

Letting P be at the origin O of our coordinate system, we write $V(P) = V$, $OM = M$, $V(M) = M$ and rewrite equation (1) as

$$M = V - \Omega \wedge M \qquad (2)$$

If we assume that V and Ω are known functions of time, then equation (2) appears as a first-order linear differential equation in M. No closed form solution, in general, exists for this equation, except when Ω is not a function of time (motion with constant angular velocity), in which case the solution is given by

$$M(t) = e^{(t-t_0)\overline{\Omega}} M(t_0) + \int_{t_0}^{t} e^{(t-s)\overline{\Omega}} V(s) ds \qquad (3)$$

where $\overline{\Omega}$ is the antisymmetric matrix representing the cross-product with Ω ($\overline{\Omega}x = \Omega$

$\wedge x$). $e^{(t-t_0)\bar{\Omega}}$ and $e^{(t-s)\bar{\Omega}}$ are rotation matrices.

We can obtain slightly more detailed results that are useful in practice by making assumptions about the functional form of $V(s)$. For example, we may assume that it is a polynomial in s:

$$V(s) = \sum_{i=0}^{n} V_i s^i$$

Using Rodrigues' equation we can write

$$e^{(t-s)\bar{\Omega}} = I + \frac{\sin((t-s)\|\Omega\|)}{\|\Omega\|}\bar{\Omega} + \frac{1-\cos((t-s)\|\Omega\|)}{\|\Omega\|^2}\bar{\Omega}^2$$

From this it is clear that, in order to compute the integral $\int_{t_0}^{t} e^{(t-s)\bar{\Omega}} V(s)\, ds$, which appears in equation (3), we need to compute the integrals

$$L_1 = \int_{t_0}^{t} s^i \sin((t-s)\|\Omega\|)\, ds$$

$$M_1 = \int_{t_0}^{t} s^i \cos((t-s)\|\Omega\|)\, ds$$

It is simple to show that can be done in closed form. For the special cases $n = 0, 1$ (constant velocity and constant acceleration) we have the following result:

proposition

when $V(s) = V + sA$, the trajectory of the point M is given by

$$M(t) = U_0 M(t_0) - U_1 V + U_2 A \qquad (4)$$

with the following values for the matrices U_i, $i = 0, 1, 2$

$$U_0 = I + \frac{\sin((t-t_0)\|\Omega\|)}{\|\Omega\|}\bar{\Omega} + \frac{1-\cos((t-t_0)\|\Omega\|)}{\|\Omega\|^2}\bar{\Omega}^2$$

$$U_1 = I(t-t_0) = \frac{1-\cos((t-t_0)\|\Omega\|)}{\|\Omega\|^2}\bar{\Omega} + \frac{(t-t_0)\|\Omega\| - \sin((t-t_0)\|\Omega\|)}{\|\Omega\|^3}\bar{\Omega}^2$$

$$U_2 = I\frac{(t-t_0)^2}{2} + \frac{(t-t_0)\|\Omega\| - \sin((t-t_0)\|\Omega\|)}{\|\Omega\|^3}\bar{\Omega} + \frac{((t-t_0)\|\Omega\|)^2 - 2(1-\cos((t-t_0)\|\Omega\|))}{2\|\Omega\|^4}\bar{\Omega}^2$$

2. CHOOSING A REPRESENTATION OF THREE-DIMENSIONAL LINE SEGMENTS

we know a representation of 3-D lines by four numbers (a, b, p and q) such that the equations of the line are ($x = az + p$, $y = bz - q$) in the first map φ_1, with the two other maps obtained by exchanging the roles played by x, y, and z. In φ_1 lines perpendicular to the z-axis cannot be represented, while in φ_2 and φ_3 it is the lines perpendicular to the x-and y-axis, respectively, that cannot be represented. From this representation, it is not too difficult to compute the weight (covariance) matrix of the endpoints M_1 and M_2 of a line segment, a 6×6 matrix.

One possible representation of a line segment is therefore the six-dimensional vector $[M_1^T, M_2^T]^T$. Just as in the two-dimensional case, this assumes that segments are oriented. If they are not, we can use the representation of direction, midpoint and length which is the vector $[a, b, M^T, l]^T$, where M is the representation of the midpoint and l is the length of the segment. Of course, weight (covariance) matrices for these representations can be computed up to first order. In what follows, we will denote by r the representation vector. and we will let C be the corresponding covariance matrix.

3. THE PLANT AND MEASUREMENT EQUATIONS

Contrary to the two-dimensional case, the state of the segment is directly related to the kinematic screw of the solid to which we assume it is attached. we define it to be the vector

$$a = \begin{bmatrix} \Omega \\ V \\ V(l) \\ \vdots \\ V(n) \end{bmatrix} \qquad (5)$$

which is a $3(n-2)$-dimensional vector in which V represents the ith order time derivative of V. In practice, n = 1 or 0. The plant equation is then

$$a_i = \Phi_{i,i-1}a_{i-1} - W_i$$

This assumes that the angular velocity is constant. We could use more elaborate models, but we would lose the closed-form expressions of proposition

Matrix $\Phi_{i,i-1}$, in the case where n = 1, is given by

$$\Phi_{i,i-1} = \begin{bmatrix} I_3 & O_3 & O_3 \\ O_3 & I_3 & I_3\Delta t \\ O_3 & O_3 & I_3 \end{bmatrix}$$

A measurement is the identification of a segment S_1 at time t_{i-1} with a segment S_2 at time t_i. Assuming that S_1 is represented by r_1 and S_2 by r_2, the measurement vector is $x = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix}$ and the measurement equation is

$$f_i(x, a) \equiv D_{i,i-1}r_1 - r_2 \qquad i \geqslant 1$$

where $D_{i,i-1}$ represents the effect on the representation vector r_1 of the rigid motion between t_{i-1} and t_i and is a function of the state a, which we will now describe.

3.1 Transforming the direction

Let u_1 be the dircetion of S_1 and suppose, for simplicity, but without loss of generality, that it can be represented in the map φ_1 in which u_2, the direction of S_2, can also be represented. Accorking to our constant angular velocity model, the direction of S_1 at time t_i will be U_0u_1, and we must have

$$\varphi_1(U_0u_1) = \varphi_1(u_2) \qquad i = 1,2,3 \qquad (6)$$

More precisely, suppose that $i=1$. If (a_1, b_1) is the representation of the direction of S_1 and (a_2, b_2) is that for S_2, we have

$$U_0 \begin{bmatrix} a_1 \\ b_1 \\ 1 \end{bmatrix} \wedge \begin{bmatrix} a_2 \\ b_2 \\ 1 \end{bmatrix} = 0 \qquad (7)$$

which expresses the fact that the directions are the same. Note that this is equivalent to two scalar equations. This is the first piece of the measurement equation, which concerns the directions.

3.2 Transforming the midpoint

Let M_1 the midpoint of S_1. In the constant acceleration model, at time t_1 it becomes $U_0M_1 + U_1V + U_2A$, and we must have

$$U_0M_1 + U_1V + U_2A = M_2 \qquad (8)$$

This is the second piece of the measurement equation, which concerns the midpoints.

3.3 Transforming the length

The length of the segment is invariant, and therefore the measurement equation is

$$l_1 = l_2 \qquad (9)$$

4. RUNNING THE PROCESS

We are now ready to put all this together in order to track a line segment in a sequence of three-dimensional frames. The situation is some what similar to that of section 3 except for the fact the measurement equations (6)—(9) are nonlinear.

4.1 Initialization

At time 0, let us consider a segment S represented by the vector r_0 with covariance matrix C_0. Let b_0 be our initial estimate of its state, and let M_0 be its weight matrix. If we have no a priori information. we assume that $b_0 = 0$, i.e., we consider that S is not moving and that M_0 diagonal. If S_i is ith segment at time 1 represented by (r_1^i, C_1^i), we consider the "possible" measurement equations

$$f_0^i(x_0^i, b_0) = 0$$

where $x_0^i = [r_0^T, r_1^{iT}]^T$ has weight matrix

$$R_0^i = \begin{pmatrix} C_0 & O \\ O & C_1^i \end{pmatrix}$$

From equations (6)—(8) we can compute the covariance matrices as follows:

$$\Lambda_0^i = \frac{\partial f_0^i}{\partial a} M_0 \frac{\partial f_0^{iT}}{\partial a} + \frac{\partial f_0^i}{\partial x} R_0^i \frac{\partial f_0^{iT}}{\partial x}$$

where each partial derivative is evaluated at (x_0^i, b_0). We then compute the Mahalanobis distances

$$d_0^i = f_0^i(x_0^i, b_0)^T (\Lambda_0^i)^{-1} f_0^i(x_0^i, b_0)$$

for all segments in the 3-D frame at time 1. Those segments with distances smaller than a fixed threshold are kept as matches.

Each match defines a token, and we update the state as follows:

$$\hat{a}_0^i = b_0 + K_0^i(x_0^i - H_0^i b_0)$$

in which

$$H_0^i = \frac{\partial f_0^i}{\partial a}(x_0^i, b_0)$$

and

$$K_0^i = M_0 H_0^i (H_0^i M_0 H_0^{iT} + R_0^i)^{-1}$$

The whight on $\hat{a}_0^i$ is the matrix $P_0^i = (M_0^{-1} + H_0^{iT}(R_0^i)^{-1} H_0^i)^{-1}$

4.2 Continuous processing

Just as in the two-dimensional case, we do the reasoning at time 2, but the generalization to an arbitrary stage follows. Let S be a token at time 1 represented by (r_1, C_1) with state $\hat{a}_0$ and with weight P_0. We make a prediction by computing the state $\hat{a}_1' = \Phi_{1,0}\hat{a}_0$ and its weight P_1'. We then determine the candidate segments at time 2. If S_i is the ith segment at time 2 represented by (r_2^i, C_2^i), we consider the "possible" measurement equations

$$f_1^i(x_1^i, \hat{a}_1') = 0$$

where $x_1^i = [r_1^T, r_2^{iT}]^T$ has weight matrix

$$R_1^i = \begin{pmatrix} C_1 & O \\ O & C_2^i \end{pmatrix}$$

and $\hat{a}_i$ has weight matrix P_1' Just as in the previous section, we select matches based on the Mahalanobis distance and update the state. We have

$$\hat{a}_1^i = \hat{a}_1' + K_1^i(x_1^i - H_1^i \hat{a}_1')$$

in which

$$H_1^i = \frac{\partial f_1^i}{\partial a}(x_1^i, \hat{a}_1')$$

and

$$K_1^i = P_1' H_1^{iT} (H_1^i P_1' H_1^{iT} + R_1^i)^{-1}$$

The weight of the new state $\hat{a}_1^i$ is the matrix $P_1^i = P_1' - K_1^i H_1' P_1'$.

5. CONCLUSION

The problem of tracking tokens in sequences of images or in sequences of stereo frames has received considerable attention in the last few years, and the use of the Kalman or extended Kalman filters or equivalently of recursive least-squares estimation theory has now become standard. The applications of these methods to the 2D-2D and 3D-3D tracking problems described in sections 3 and 4 have not been as numerous as their applications to the 2D-3D tracking problem in which the observations are made in the image and the tracking is done in three dimensions.

6. REFERENCES

H. Shariat and K. E. Price. Motion estimation with more than two frames. IEEE Trans. PAMI, 12(5): 417—434, May 1990.

Z. Zhang. Motion Analysis from a Sequence of Stereo Frames and Its Applications. Ph. D. thesis, 1990.

Z. Zhang and O. D. Faugeras. 3D Dynamic Scene Analysis: A Stereo Based Approach. Springer-Verlag, 1992.