COMPUTER VISION, GRAPHICS, AND IMAGE PROCESSING 34, 372–387 (1986)

NOTE

A Note on "Rigid Body Motion from Depth and Optical Flow"

ROBERT M. HARALICK

AND

XINHUA ZHUANG*

Machine Vision International, Ann Arbor, Michigan 48104

Received May 31, 1984; revised February 8, 1986

A theoretical solution to the problem "Rigid Body Motion from Depth and Optical Flow" is uniquely determined, given four uncoplanar initial spatial points and, associated with them, optical flow and depth information. A numerical example is given to explain the theory and related algorithm.

I. INTRODUCTION

In [1] Ballard and Kimball consider the problem "rigid body motion from depth and optical flow." They argue that motion parameters are computable in parallel from depth and optical flow information and, when coupled to the flow computations, the rigid body computations can resolve difficult singularities in the flow calculations. Although many insights which the authors propose are useful, the method they suggest to solve motion parameters from depth and optical flow information is not correct. In Section II, following their basic conventions, an analysis which points out the mistake in [1] is given.

In Section III, we solve the rigid body motion from depth and optic flow problem by a differential equation approach. It is shown that if at any time, given depth and optic flow of four uncoplanar object points which might not be the same four points from time to time, then the rigid body motion is uniquely determined. If these four uncoplanar object points are not changable from time to time, then there is a straightforward approach to recover the motion as seen from the argument in the section. Finally, a numeric example is used to explain the theory and related algorithm.

In summary, the problem, "rigid body motion from depth and optical flow" is solved without the assumption that the body is not subject to large external forces, which is required by the authors in [1] and also without the assumption that the depth and optic flow come from the same object points from time to time.

II. BALLARD AND KIMBALL'S METHOD

The perspective transform describes how points in 3-dimensional space are projected onto the retina. Figure 1 shows an imaging system with the lens located at

*Visiting Scientist from Zhejiang Computing Technique Institute, Hangzhou, People's Republic of China.

0734-189X/86 $3.00

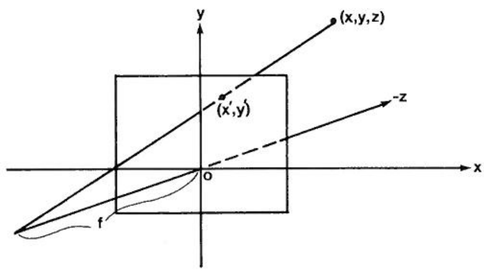

FIG. 1. Viewing geometry.

position f on the z axis and the image focal plane or retina is at $z = 0$. Points with negative z values are imaged.

Referring to the figure, the retinal position $X' = (x', y')^T$ is related to a given 3-dimensional point $X = (x, y, z)^T$ by

$$x' = xf/(f - z)$$
$$y' = yf/(f - z) \tag{1}$$

where f is the optical focal length. Differentiating these equations with respect to time results in a relationship between optical flow and what we call 3D flow. The optical flow describes the field of instantaneous retinal velocities of a group of retinal points $\{X'\}$. Optical flow is denoted by

$$\frac{dX'}{dt} = \begin{pmatrix} \dfrac{dx'}{dt} \\ \dfrac{dy'}{dt} \end{pmatrix} = \begin{pmatrix} u(x'(t), y'(t), t) \\ v(x'(t), y'(t), t) \end{pmatrix}$$

and 3D flow by

$$\frac{dX}{dt} = \begin{pmatrix} \dfrac{dx}{dt} \\ \dfrac{dy}{dt} \\ \dfrac{dz}{dt} \end{pmatrix} = \begin{pmatrix} V_x(x(t), y(t), z(t), t) \\ V_y(x(t), y(t), z(t), t) \\ V_z(x(t), y(t), z(t), t) \end{pmatrix}.$$

The following equations hold as [1] indicates:

$$(f - z)u = v_x f + x'v_z$$
$$(f - z)v = v_y f + y'v_z. \tag{2}$$

Assume that depth information $z(x'(t), y'(t), t)$ and optical flow $(u(x'(t), y'(t), t), v(x'(t), y'(t), t))$ are known. Then, the 3D position information is $X(x', y', t) = (x(x'(t), y'(t), t), y(x'(t), y'(t), t), z(x'(t), y'(t), t))^{\mathrm{T}}$, and 3D flow are known by

$$X(x', y', t) = \left(\frac{x'(f - z(x', y', t))}{f}, \frac{y'(f - z(x', y', t))}{f}, z(x', y', t) \right)^{\mathrm{T}} \quad (3)$$

and letting $z_{x'}$ denote the partial of z with respect to x', $z_{y'}$ denote the partial of z with respect to y', and z_t denote the partial of z with respect to t,

$$\begin{aligned} V &= (v_x, v_y, v_z)^{\mathrm{T}} \\ &= \left(\frac{(f - z)u - x'(z_{x'}u + z_{y'}v + z_t)}{f}, \right. \\ &\qquad \left. \frac{(f - z)v - y'(z_{x'}u + z_{y'}v + z_t)}{f}, z_{x'}u + z_{y'}v + z_t \right)^{\mathrm{T}} \end{aligned} \quad (4)$$

as the authors in [1] demonstrate.

Let b be a fixed point with respect to the rigid body motion. There is a fixed coordinate frame F centered at b with respect to the rigid body. The motion of an arbitrary point X with respect to the world coordinate frame can be represented as the sum of the motion $x_b(t)$ of the fixed point b plus the rotation of X around and axis through b.

$$X(t) = X_b(t) + \rho(t) = X_b(t) + R^*(t)\rho(0) \quad (5)$$

where $R^*(t)$ is an orthogonal matrix with $\det(R^*(t)) = 1$; $\rho(t)$ represents the vector $X(t)$ in F at the time t; $X(t)$ represents the point X on the rigid body at the time t in the world coordinate frame.

An alternative way is to describe the rigid body motion as the sum of the translation plus the rotation around an axis through the origin 0;

$$X(t) = R(t)X(0) + T(t). \quad (5')$$

Especially,

$$X_b(t) = R(t)X_b(0) + T(t) \quad (5'')$$

Combining (5), (5′), (5″) together leads to $R^*(t) = R(t)$ and $T(t) = X_b(t) - R(t)X_b(0)$. Think about the instantaneous rigid body motion from the time t to the time $t + \Delta t$ with Δt small enough. Then, the rigid body experiences two kinds of motion: a small translation along the orientation of $V_b(t)$, $X_b(t + t\Delta) - X_b(t) \doteq V_b(t)\,\Delta t$ and a small rotation around some axis through b with an orientation $\omega(t)(|\omega| = 1)$ by a small angle $\Delta\theta$, $\rho(t + \Delta t) - \rho(t) = \Delta\theta\omega(t) \times \rho(t)$ (Remember a small rotation could be approximated by a cross product!). Thus,

$$\frac{d\rho(t)}{dt} = \left[\frac{d\theta}{dt}\omega(t)\right] \times \rho(t).$$

Hence, $\Omega(t)$ defined by $[(d\theta/dt)\,\omega(t)]$ represents the instantaneous angular velocity of rotation around the axis through b with the orientation $\omega(t)$. As a result,

$$V(t) = V_b(t) + \dot{\rho}(t) = V_b(t) + \Omega(t) \times \rho(t). \tag{6}$$

Another expression for $V(t)$ comes from (5′):

$$V(t) = \dot{R}(t)X(0) + \dot{T}(t). \tag{6′}$$

It could be easily verified that

$$V_b(t) = \dot{R}(t)X_b(0) + \dot{T}(t) \tag{7}$$

$$\dot{R}(t)R^{\mathrm{T}}(t)\rho(t) = \Omega(t) \times \rho(t). \tag{8}$$

Noticing $R(t)R^{\mathrm{T}}(t) = I_3$ and hence $\dot{R}(t)R^{\mathrm{T}}(t) = -(\dot{R}(t)R^{\mathrm{T}}(t))^{\mathrm{T}}$, we realize that the matrix $(\dot{R}(t)R^{\mathrm{T}}(t))$ is skew-symmetric. It is easy to directly verify by using (8).

$$\Omega(t) = -\begin{pmatrix} \langle \dot{r}_2, r_3 \rangle \\ \langle \dot{r}_3, r_1 \rangle \\ \langle \dot{r}_1, r_2 \rangle \end{pmatrix} \begin{matrix} = \\ = \\ = \end{matrix} \begin{pmatrix} \langle \dot{r}_3, r_2 \rangle \\ \langle \dot{r}_1, r_3 \rangle \\ \langle \dot{r}_2, r_1 \rangle \end{pmatrix} \tag{9}$$

(see (6′)) where

$$R = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix}, \quad \text{and} \quad \langle \cdot\,, \cdot \rangle$$

is the scalar product between two vectors. Furthermore, we could obtain the expression for acceleration $a(t)$:

$$a(t) = a_b(t) + (d\Omega(t)/dt) \times \rho(t) + (\Omega(t) \cdot \rho(t))\Omega(t) - |\Omega(t)|^2\rho(t). \tag{10}$$

Equations (5), (6), and (10) are established in [1], too, where $\Omega(t) \cdot \rho(t)$ is the scalar product between the two column vectors $\Omega(t)$ and $\rho(t)$.

The problem discussed in [1] is: given depth information $z(x', y', t)$ and optical flow $(x', y';\ u, v)$, determine a procedure which computes the motion parameters $X_b(t)$, $V_b(t)$, and $\Omega(t)$. Given depth information $z_b(t)$ and optical flow $(x_b'(t), y_b'(t);\ u_b(t), v_b(t))$ which correspond to $X_b(t)$, both $X_b(t)$ and $V_b(t)$ can be calculated by (3) and (4) as authors in [1] show. In this case, only $\Omega(t)$ needs to be determined. There they assume that

$$a_b(t) \doteq 0 \tag{11}$$

and

$$\frac{d\Omega(t)}{dt} \doteq 0. \tag{11'}$$

Then they correctly show

$$a(t)\Omega(t) \doteq 0. \tag{12}$$

Based on the approximate equality (12) they proceed as follows: let

$$\begin{aligned} a_1 &= a(t + \Delta t), \\ a_2 &= a(t + 2\Delta t). \end{aligned} \tag{13}$$

Then they argue

$$a_1 \doteq a, \qquad a_2 \doteq a. \tag{14}$$

Thus, by (12) we have

$$a_1 \cdot \Omega \doteq 0, \qquad a_2 \cdot \Omega \doteq 0, \tag{15}$$

that is, Ω is approximately perpendicular to both a_1 and a_2. It is always true that

$$|\Omega \cdot (a_1 \times a_2)|^2 + |\Omega \times (a_1 \times a_2)|^2 = |\Omega|^2 |a_1 \times a_2|^2 \tag{16}$$

and

$$\Omega \times (a_1 \times a_2) = a_1(\Omega \cdot a_2) - a_2(\Omega \cdot a_1). \tag{17}$$

It is obvious that Ω is parallel to $a_1 \times a_2$ iff $\Omega \times (a_1 \times a_2) = 0$. Hence if $\Omega \cdot a_2 \doteq 0$ and $\Omega \cdot a_1 \doteq 0$, they conclude that $|\Omega \cdot (a_1 \times a_2)|^2 \doteq |\Omega|^2 |a_1 \times a_2|^2$ which would make Ω parallel to $a_1 \times a_2$. Unfortunately, this reasoning is not correct. Some care has to be taken due to the fact that the approximations are only first order. Carefully tracking the approximation under the exact assumption $a \cdot \Omega = 0$ we easily obtain not only

$$\begin{aligned} a_1 \cdot \Omega &\doteq \Delta t \dot{a} \cdot \Omega \\ a_2 \cdot \Omega &\doteq 2\Delta t \dot{a} \cdot \Omega \end{aligned} \tag{18}$$

but also

$$a_1 \times a_2 \doteq \Delta t a \times \dot{a}. \tag{19}$$

Equation (18) means Ω is perpendicular to both a_1 and a_2 to within a first-order infinitesimal. Thus, Ω is parallel to $a_1 \times a_2$ to within a first-order infinitesimal (see the identity (17)). However, (19) does mean that $a_1 \times a_2$ itself is also a first-order infinitesimal. In this case, it is hard to say whether or not Ω is approximately parallel to $a_1 \times a_2$. To explain this error concretely we work out the correct answers for the special case

$$\frac{d^2\Omega}{dt^2} \doteq 0 \tag{20}$$

and show that in this case that it is necessary to have $\dot{a}_b \doteq 0$ in order for Ω to be parallel to $a_1 \times a_2$. From (10), (11), (11′),

$$\begin{aligned} a(t) &\doteq (\Omega(t)\cdot\rho(t))\Omega(t) - |\Omega(t)|^2\rho(t) \\ \dot{a}(t) &\doteq \dot{a}_b(t) + (\Omega(t)\cdot\dot{\rho}(t))\Omega(t) - |\Omega(t)|^2\dot{\rho}(t). \end{aligned} \tag{21}$$

But $\dot{\rho}(t) = \Omega(t) \times \rho(t)$. Hence

$$\begin{aligned} a(t) &\doteq \dot{a}_b(t) + (\Omega(t)\cdot[\Omega(t)\times\rho(t)])\Omega(t) - |\Omega(t)|^2[\Omega(t)\times\rho(t)] \\ &= \dot{a}_b - |\Omega|^2[\Omega\times\rho] \end{aligned} \tag{22}$$

$$\begin{aligned} a\times\dot{a} &\doteq a\times\dot{a}_b - \left[(\Omega\cdot\rho)\Omega - |\Omega|^2\rho\right]\times|\Omega|^2[\Omega\times\rho] \\ &= a\times\dot{a}_b + |\Omega|^4\rho\times[\Omega\times\rho] - |\Omega|^3|\rho|\cos\theta\,\Omega\times[\Omega\times\rho]. \end{aligned} \tag{23}$$

It is easy to verify

$$\begin{aligned} \rho\times[\Omega\times\rho] &= |\Omega||\rho|^2\sin\theta\frac{|\rho|\Omega - |\Omega|\cos\theta\rho}{\sqrt{|\rho|^2|\Omega|^2 - (\rho\Omega)^2}} \\ &= |\rho|(|\rho|\Omega - |\Omega|\cos\theta\rho) \end{aligned} \tag{24}$$

$$\begin{aligned} \Omega\times[\Omega\times\rho] &= |\Omega|^2|\rho|\sin\theta\frac{|\rho|\cos\theta\Omega - |\Omega|\rho}{\sqrt{|\rho|^2|\Omega|^2 - (\rho\Omega)^2}} \\ &= |\Omega|(|\rho|\cos\theta\Omega - |\Omega|\rho) \end{aligned} \tag{25}$$

where θ is the angle between Ω and ρ, $0 \le \theta \le \pi$. Thus, we obtain

$$\begin{aligned} a\times\dot{a} &\doteq a\times\dot{a}_b + |\Omega|^4|\rho|\{|\rho|\Omega - |\Omega|\cos\theta\rho - |\rho|\cos^2\theta\Omega + |\Omega|\cos\theta\rho\} \\ &= a\times\dot{a}_b + |\Omega|^4|\rho|^2\sin^2\theta\Omega \\ &= (\Omega\cdot\rho)(\Omega\times\dot{a}_b) - |\Omega|^2(\rho\times\dot{a}_b) + |\Omega|^4|\rho|^2\sin^2\theta\Omega. \end{aligned} \tag{26}$$

Combining (26) with $a_1 \times a_2 \doteq a \times \dot{a}(\Delta t)$, we conclude that for arbitrary choice of ρ, Ω can be approximately parallel to $a_1 \times a_2$ only if

$$\dot{a}_b \doteq 0. \tag{27}$$

Otherwise, the vector

$$\begin{aligned}(\Omega\rho)(\Omega \times \dot{a}_b) - |\Omega|^2(\rho \times \dot{a}_b) &= |\rho|\{(\Omega \cdot \rho/|\rho|)(\Omega \times \dot{a}_b) \\ &\quad -|\Omega|^2((\rho/|\rho|) \times \dot{a}_b)\} \\ &= |\rho|\{|\Omega|\cos\theta(\Omega \times \dot{a}_b) - |\Omega|^2((\rho/|\rho|) \times \dot{a}_b)\} \quad (28)\end{aligned}$$

could have various orientations so that a × ȧ has significantly different orientations from Ω's orientation. It should be pointed out that the condition (27) is not implied by the small external forces assumption: $a_b \doteq 0$ and $\dot{\Omega} \doteq 0$. Even when the condition (27) is added in, concluding that $a_1 \times a_2$ is parallel to Ω is not always reliable. When ρ is nearly parallel to the rotation vector Ω (in this case $\theta \doteq 0$), $a_1 \times a_2$ would produce an estimation of Ω's orientation with large deviation since it is probable to have

$$|a \times \dot{a} - |\Omega|^4|\rho|^2\sin^2\theta\Omega| > ||\Omega|^4|\rho|^2\sin^2\theta\Omega|. \quad (29)$$

Therefore, when a, a_1, a_2 all come from the same rigid body motion, the idea that $\Omega/|\Omega|$ should be the same as $(a_1 \times a_2)/(|a_1 \times a_2|)$ within the accuracy of the measurement is not correct in general.

EXAMPLE. The rigid body motion is determined by

$$\begin{aligned}X_b(t) &= -\varepsilon \sin t\Omega_0, \qquad \varepsilon > 0 \text{ small}, \\ V_b(t) &= -\varepsilon \cos t\Omega_0, \\ \Omega(t) &= \Omega_0, \qquad |\Omega_0| = 1, \\ t &\geq 0. \end{aligned} \quad (30)$$

From (29), we have

$$a_b(t) = \varepsilon \sin t\Omega_0, \qquad a_b(0) = 0, \quad (31)$$

$$\dot{a}_b(t) = \varepsilon \cos t\Omega_0, \qquad \dot{a}_b(0) = \varepsilon\Omega_0 \quad (32)$$

$$\dot{\Omega}(t) = \ddot{\Omega}(t) = 0 \quad (33)$$

$$a(0) = (\Omega_0 \cdot \rho(0))\Omega_0 - |\Omega_0|^2\rho(0) \quad (34)$$

$$\dot{a}(0) = \varepsilon\Omega_0 - |\Omega_0|^2[\Omega_0 \times \rho(0)] \quad (35)$$

$$\begin{aligned}a(0) \times \dot{a}(0) &= (\Omega_0 \cdot \rho(0))(\Omega_0 \times \varepsilon\Omega_0) - |\Omega_0|^2(\rho(0) \times \varepsilon\Omega_0) \\ &\quad +|\Omega_0|^4|\rho(0)|^2 \sin^2\theta\Omega_0. \\ &= -\varepsilon|\Omega_0|^2(\rho(0) \times \Omega_0) + |\Omega_0|^4|\rho(0)|^2 \sin^2\theta\Omega_0 \\ &= -\varepsilon(\rho(0) \times \Omega_0) + |\rho(0)|^2 \sin^2\theta\Omega_0. \end{aligned} \quad (36)$$

The two vectors $\rho(0) \times \Omega_0$ and Ω_0 are perpendicular to each other. And

$$\left|-\varepsilon(\rho(0) \times \Omega_0)\right| = \varepsilon|\rho(0)|\sin\theta \tag{37}$$

$$\left||\rho(0)|^2 \sin^2\theta\Omega_0\right| = |\rho(0)|^2 \sin^2\theta. \tag{38}$$

When $0 < \theta < \varepsilon$ and $|\rho(0)| = 1$, we obtain

$$\left|-\varepsilon(\rho(0) \times \Omega_0)\right| > \left||\rho(0)|^2 \sin^2\theta\Omega_0\right|. \tag{39}$$

Therefore, in this case the angle α between $a(0) \times \dot{a}(0)$ and Ω should be in the interval $(\pi/4, \pi/2)$. For this kind $\rho(0)$, $a(0) \times \dot{a}(0)$ is absolutely not parallel to Ω_0. Furthermore

$$\begin{aligned} a_1(0) \times a_2(0) &= a(\Delta t) \times a(2\Delta t) \\ &\doteq \Delta t(a(0) \times \dot{a}(0)) \end{aligned} \tag{40}$$

implies that for any fixed $\rho(0), (a_1 \times a_2)$'s orientation approaches the orientation of $(a(0) \times \dot{a}(0))$ when $\Delta t \to 0$. Therefore, for $\rho(0)$ with $|\rho(0)| = 1, 0 < \theta < \varepsilon$ and Δt small enough, since the angle between $a(0) \times \dot{a}(0)$ and Ω_0 is in the interval $(\pi/4, \pi/2)$, $a_1 \times a_2$ is absolutely not parallel to Ω_0, too.

In summary, the previous analysis and semi-numeric example indicate that the idea considering $\Omega/|\Omega|$ being the same as $(a_1 \times a_2)/(|a_1 \times a_2|)$ within the accuracy of the measurement is not correct in general even under the stronger assumptions: $a_b, \dot{a}_b, \dot{\Omega}, \ddot{\Omega} \doteq 0$.

In the following section, we develop a general procedure to solve the problem "Rigid Body Motion from Depth and Optical Flow." The procedure does not use any assumptions like the small external forces and $X_b(0)$, $V_b(t)$ known.

III. A DIFFERENTIAL EQUATION APPROACH FOR SOLVING THE RIGID BODY MOTION FROM DEPTH AND OPTICAL FLOW

In this section, we use the world coordinate system as shown in Fig. 1. For convenience, we change the notational conventions a little bit. Let $P = (x, y, z)^{\mathrm{T}}$ and $p' = (x', y')^{\mathrm{T}}$. Assume that $\mathbb{P}'(t)$ is a set of retinal points $\{p'(t)\}$ and $\mathbb{P}(t) = \{P(p'(t), t)\}$ is a corresponding set of object points determined from $\mathbb{P}'(t)$ and the depth information $z(p'(t), t)$, where $\mathbb{P}(t)$ might not be the same object points from time to time. The rigid body motion is represented as

$$P(t) = R(t)P(0) + T(t) \tag{41}$$

where $P(t)$ and $P(0)$ represent spatial coordinates of the same object point on the rigid body at times t and 0, respectively. $R(t)$ is the rotation matrix and $T(t)$ the translation vector. According to the convention, $R(0) = I$ (a 3×3 identity matrix) and $T(0) = 0$. As known, $R(t)$ is an orthonormal matrix $(R(t)R^{\mathrm{T}}(t) = I)$ with $\det(R(t)) = 1$.

Assume that for each $p'(t) \in \mathbb{P}'(t)$ the depth information $z(p'(t), t)$ and the optical flow $(u(p'(t), t), v(p'(t), t))$ are given. Then, $P(p'(t), t)$ and $\dot{P}(p'(t), t)$ (denoted by $w(p'(t), t)$) are determined respectively by

$$P(x', y', t) = \left(\frac{x'(f - z(x', y', t))}{f}, \frac{y'(f - z(x', y', t))}{f}, z(x', y', t) \right)^{\mathrm{T}} \tag{42}$$

$$\begin{aligned} w(p'(t), t) &= (w_1(p'(t), t), w_2(p'(t), t), w_3(p'(t), t))^{\mathrm{T}} \\ &= \left(\frac{(f - z)u - x'(z_{x'}u + z_{y'}v + z_t)}{f}, \right. \\ &\qquad \frac{(f - z)v - y'(z_{x'}u + z_{y'}v + z_t)}{f}, \\ &\qquad \left. z_{x'}u + z_{y'}v + z_t \right)^{\mathrm{T}}. \end{aligned} \tag{43}$$

To determine the motion parameters $R(t)$ and $T(t)$, we only need to establish the differential equations which govern them since the initial values $R(0) = I$ and $T(0) = 0$ are known in advance. Consider

$$P(t + \Delta t) = R(t + \Delta t)P(0) + T(t + \Delta t) \tag{44}$$

$$= G(t, \Delta t)P(t) + H(t, \Delta t) \tag{45}$$

where the rotation $G(t, \Delta t)$ and the translation $H(t, \Delta t)$ represent the rigid body motion from the time t to the time $t + \Delta t$. Substituting $R(t)P(0) + T(t)$ for $P(t)$ in (45) results in

$$\begin{aligned} &\{R(t + \Delta t) - G(t, \Delta t)R(t)\}P(0) \\ &\quad + \{T(t + \Delta t) - G(t, \Delta t)T(t) - H(t, \Delta t)\} = 0. \end{aligned} \tag{46}$$

Since $P(0)$ comes from the rigid body with nonzero volume, (46) implies

$$R(t + \Delta t) = G(t, \Delta t)R(t) \tag{47}$$

and

$$T(t + \Delta t) = G(t, \Delta t)T(t) + H(t, \Delta t). \tag{48}$$

Thus

$$G(t, \Delta t) = R(t + \Delta t)R^{\mathrm{T}}(t) \tag{49}$$

and

$$H(t, \Delta t) = T(t + \Delta t) - R(t + \Delta t)R^{T}(t)T(t). \tag{50}$$

Noticing

$$R(t + \Delta t) \doteq R(t) + \dot{R}(t)\,\Delta t \tag{51}$$

and

$$T(t + \Delta t) \doteq T(t) + \dot{T}(t)\,\Delta t \tag{52}$$

we have

$$G(t, \Delta t) \doteq I + \dot{R}(t)R^{T}(t)\,\Delta t \tag{53}$$

and

$$H(t, \Delta t) \doteq \left(\dot{T}(t) - \dot{R}(t)R^{T}(t)T(t)\right)\Delta t \tag{54}$$

Thus, combining (53), (54) with (45), we obtain

$$\frac{P(t + \Delta t) - P(t)}{\Delta t} \doteq \dot{R}(t)R^{T}(t)P(t) + \left(\dot{T}(t) - \dot{R}(t)R^{T}(t)T(t)\right). \tag{55}$$

Let

$$S(t) = \dot{R}(t)R^{T}(t) \tag{56}$$

$$K(t) = \dot{T}(t) - S(t)T(t). \tag{57}$$

Letting Δt approach to zero, finally we obtain

$$w(t) = S(t)P(t) + K(t), \tag{58}$$

where $S(t)$ and $K(t)$ actually represent the instantaneous description of motion.

Assume that $\mathbb{P}'(T)$ contains points $\{p_i'(t);\ i = 1, 2, 3, 4\}$ so that the corresponding four object points $\{P_i(t) = P(p_i'(t), t);\ i = 1, 2, 3, 4\} \subset \mathbb{P}(t)$ are not coplanar. Denote the corresponding velocities $w(p_i'(t), t)$ by $w_i(t)$. $P_i(t)$'s and $w_i(t)$'s are uniquely determined from the optical flow information $\{u_i(t) = u(p_i'(t), t), v_i(t) = v(p_i'(t), t);\ i = 1, 2, 3, 4\}$ and the depth information $\{z_i(t) = z(p_i'(t), t);\ i = 1, 2, 3, 4\}$ by using formulas (42) and (43). Once w_i's and P_i's are determined, $S(t)$ and $K(t)$ can be uniquely determined as

$$S(t) = [w_2 - w_1, w_3 - w_1, w_4 - w_1][P_2 - P_1, P_3 - P_1, P_4 - P_1]^{-1}, \tag{59}$$

$$K(t) = w_1 - S(t)P_1. \tag{60}$$

Here the inverse matrix $[P_2 - P_1, P_3 - P_1, P_4 - P_1]^{-1}$ exists since the vectors P_1, P_2, P_3, P_4 are not coplanar. Hence, it is clear that the instantaneous description of motion can be simply recovered from depth and optic flow of four uncoplanar object points.

If four points P_1, P_2, P_3, P_4 are coplanar, we have to require that there are three within them, for instance P_1, P_2, P_3 such that the corresponding determinant $\det[P_1, P_2, P_3]$ is nonzero. When there are no such three points within them, the motion cannot be uniquely determined. To explain this, we need to modify the instantaneous motion equation a little bit. Let us first point out that the matrix S is skew-symmetric: $S^{\mathrm{T}} = -S$. In fact,

$$\begin{aligned} S + S^{\mathrm{T}} &= \dot{R}(t)R^{\mathrm{T}}(t) + R(t)\dot{R}^{\mathrm{T}}(t) \\ &= \frac{d}{dt}\left[R(t)R^{\mathrm{T}}(t)\right] = \dot{I} = 0. \end{aligned}$$

Second, for any skew-symmetric matrix

$$S = \begin{bmatrix} 0 & -w_3 & w_2 \\ w_3 & 0 & -w_1 \\ -w_2 & w_1 & 0 \end{bmatrix}$$

letting $\Omega = [w_1, w_2, w_3]^{\mathrm{T}}$ we could directly verify

$$SP = \Omega \times P.$$

Thus, the instantaneous motion equation could be rewritten as

$$w = \Omega \times P + K.$$

Now suppose we are given the true instantaneous motion Ω, K. Pick any plane which passes through the origin and any three nonzero vectors Q_1, Q_2, Q_3 lying on the plane. For any nonzero vector N which lies on the plane and uncolinear with any of the Q_i's there are three nonzero parameters $\lambda_1, \lambda_2, \lambda_3$ such that there holds

$$N \times (\lambda_1 Q_1) = N \times (\lambda_2 Q_2) = N \times (\lambda_3 Q_3) \neq 0.$$

Let $P_i = \lambda_i Q_i$ and $w_i = \Omega \times P_i + K$. Then, it is clear that the following equalities hold too:

$$w_i = (\Omega + N) \times P_i + (K - N \times P_1)$$

which indicates the motion cannot be uniquely determined by these three points P_1, P_2, P_3 and their velocities w_1, w_2, w_3.

For recovering the evolution of motion with t, we need to solve $R(t)$ and $T(t)$ from $S(t)$ and $K(t)$. As such, we obtain the following differential equation for $R(t)$

and $T(t)$:

$$\dot{R}(t)R^{\mathrm{T}}(t) = S(t) \tag{61}$$

$$\dot{T}(t) - S(t)T(t) = K(t) \tag{62}$$

where $S(t)$ and $K(t)$ are considered known.

It is worth pointing out that for different times, P_1, P_2, P_3, P_4 do not necessarily represent the same set of object points. If they do, then $\det([P_2 - P_1, P_3 - P_1, P_4 - P_1])$ is unchanged when the time evolves and the motion can be solved in a straightforward way.

Recalling $R^{\mathrm{T}}(t)R(t) = I$, we could transform (61) into

$$\dot{R}(t) - S(t)R(t) = 0. \tag{63}$$

It is classical in theory of differential equations that the differential equation (63) with the initial value $R(0) = I$ has the general solution

$$R(t) = \exp\left(\int_0^t S(\tau)\, d\tau\right) \tag{64}$$

and the differential equation (62) with the initial value $T(0) = 0$ has the general solution

$$\begin{aligned} T(t) &= \int_0^t \exp\left(\int_\tau^t S(\xi)\, d\xi\right) K(\tau)\, d\tau \\ &= \int_0^t \exp\left(\int_0^t S(\xi)\, d\xi - \int_0^\tau S(\xi)\, d\xi\right) K(\tau)\, d\tau \\ &= \int_0^t \exp\left(\int_0^t S(\xi)\, d\xi\right) \exp\left(-\int_0^\tau S(\xi)\, d\xi\right) K(\tau)\, d\tau \\ &= R(t) \int_0^t R^{-1}(\tau) K(\tau)\, d\tau \\ &= R(t) \int_0^t R^{\mathrm{T}}(\tau) K(\tau)\, d\tau \end{aligned} \tag{65}$$

where the integral of a matrix's function $M(t) = [M_{ij}(t)]$ is defined by

$$\int_0^t M(\tau)\, d\tau = \left[\int_0^t M_{ij}(\tau)\, d\tau\right] \tag{66}$$

and the exponential of a matrix M is defined by

$$\exp(M) = \sum_{k \geq 0} \frac{1}{k!} M^k. \tag{67}$$

We can directly verify the correctness of formula (64) and (65) as follows:

$$R(0) = \exp(0) = I,$$
$$\dot{R}(t) = \frac{d}{dt}\left(\int_0^t S(\tau)\,d\tau\right)\exp\left(\int_0^t S(\tau)\,d\tau\right)$$
$$= S(t)R(t),$$
$$T(0) = 0,$$
$$\dot{T}(t) = \dot{R}(t)\int_0^t R^T(\tau)K(\tau)\,d\tau + R(t)R^T(t)K(t)$$
$$= S(t)R(t)\int_0^t R^T(\tau)K(\tau)\,d\tau + K(t)$$
$$= S(t)T(t) + K(t).$$

As a result, four optical flow points and the corresponding depths which come from four uncoplanar object points can uniquely determine the rigid body motion. An iterative scheme to compute $R_k = R(t_k)$ and $T_k = T(t_k)$ could be directly obtained from (63) and (62) with the initial values $R(0) = I$ and $T(0) = 0$ as follows:

$$\begin{aligned}
R_0 &= I,\\
T_0 &= 0,\\
R_k &= R_{k-1} + S_{k-1}R_{k-1}\Delta_k\\
&= \{I + S_{k-1}\Delta_k\}R_{k-1},\\
T_k &= T_{k-1} + \{S_{k-1}T_{k-1} + K_{k-1}\}\Delta_k\\
&= \{I + S_{k-1}\Delta_k\}T_{k-1} + K_{k-1}\Delta_k,\\
\Delta_k &= t_k - t_{k-1},\, t_{k-1} < t_k,\, k \geq 1,
\end{aligned} \tag{68}$$

or equivalently,

$$\begin{aligned}
R_k &= \prod_{i=1}^{k}[I + S_{i-1}\Delta_i],\\
T_k &= \sum_{j=2}^{k}\left(\prod_{i=j}^{k}[I + S_{i-1}\Delta_i]\right)K_{j-2}\Delta_{j-1} + K_{k-1}\Delta_k, \qquad k \geq 1.
\end{aligned} \tag{69}$$

Since $\dot{R}(t)R^{\mathrm{T}}(t) + R(t)\dot{R}^{\mathrm{T}}(t) = 0$, $S(t)$ should be skew-symmetric: $S^T(t) = -S(t)$. Thus,

$$\begin{aligned}
R_k R_k^T &= \prod_{i=1}^{k}\left[I + S_{i-1}\Delta_i + S_{i-1}^{\mathrm{T}}\Delta_i + S_{i-1}S_{i-1}^{\mathrm{T}}(\Delta_i)^2\right]\\
&= \prod_{i=1}^{k}\left[I - S_{i-1}^2(\Delta_i)^2\right]\\
&= I - \sum_{i=1}^{k} S_{i-1}^2(\Delta_i)^2\\
&= I \qquad \text{(if all } \Delta_i\text{'s are small enough!)}
\end{aligned} \tag{70}$$

Therefore, when all Δ_k's are small enough, not only R_k and T_k approximate $R(t_k)$ and $T(t_k)$ very well, but also R_k keeps the orthogonality very well.

In summary, given the optical flow information $\{p_i'(t), [u(p_i'(t), t), v(p_i'(t), t)];\ i = 1,2,3,4\}$ and the depth information $\{z(p_i'(t), t);\ i = 1,2,3,4\}$ such that the corresponding object points $\{P(p_i'(t), t);\ i = 1,2,3,4\}$ are not coplanar, the motion parameters are uniquely determined. If the optical flow $\{p_i'(t), [u(p'(t), t), v(p_i'(t), t)];\ i = 1,2,3,4\}$ traces the motion of the same object point set, then we only need to assume the initial four object points $\{P_i(0);\ i = 1,2,3,4\}$ being uncoplanar.

EXAMPLE. Given the optical flow

$$p': \begin{vmatrix} \dfrac{\sin t - \cos t}{2} \\ \dfrac{-\sin t - \cos t}{2} \end{vmatrix}, \begin{vmatrix} \dfrac{\sin t}{2} \\ \dfrac{-\cos t}{2} \end{vmatrix}, \begin{vmatrix} \dfrac{-\cos t}{2} \\ \dfrac{-\sin t}{2} \end{vmatrix}, \begin{vmatrix} \sin t - \cos t \\ -\sin t - \cos t \end{vmatrix}$$

$$\frac{u}{v}: \begin{vmatrix} \dfrac{\sin t + \cos t}{2} \\ \dfrac{\sin t - \cos t}{2} \end{vmatrix}, \begin{vmatrix} \dfrac{\cos t}{2} \\ \dfrac{\sin t}{2} \end{vmatrix}, \begin{vmatrix} \dfrac{\sin t}{2} \\ \dfrac{-\cos t}{2} \end{vmatrix}, \begin{vmatrix} \cos t + \sin t \\ \sin t - \cos t \end{vmatrix}$$

and the depth information

$$z: -1, -1, -1, 0,$$

as well as the focal length $f = 1$, the corresponding object points and velocities are obtained by using (42) and (43):

$$P: \begin{vmatrix} \sin t - \cos t \\ -\sin t - \cos t \\ -1 \end{vmatrix}, \begin{vmatrix} \sin t \\ -\cos t \\ -1 \end{vmatrix}, \begin{vmatrix} -\cos t \\ -\sin t \\ -1 \end{vmatrix}, \begin{vmatrix} \sin t - \cos t \\ -\sin t - \cos t \\ 0 \end{vmatrix}$$

$$w: \begin{vmatrix} \sin t + \cos t \\ \sin t - \cos t \\ 0 \end{vmatrix}, \begin{vmatrix} \cos t \\ \sin t \\ 0 \end{vmatrix}, \begin{vmatrix} \sin t \\ -\cos t \\ 0 \end{vmatrix}, \begin{vmatrix} \sin t + \cos t \\ \sin t - \cos t \\ 0 \end{vmatrix}.$$

Thus

$$\begin{aligned} S(t) &= [w_2 - w_1, w_3 - w_1, w_4 - w_1][P_2 - P_1, P_3 - P_1, P_4 - P_1]^{-1} \\ &= \begin{vmatrix} -\sin t & -\cos t & 0 \\ \cos t & -\sin t & 0 \\ 0 & 0 & 0 \end{vmatrix} \begin{vmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{vmatrix}^{-1} \\ &= \begin{vmatrix} -\sin t & -\cos t & 0 \\ \cos t & -\sin t & 0 \\ 0 & 0 & 0 \end{vmatrix} \begin{vmatrix} \cos t & \sin t & 0 \\ -\sin t & \cos t & 0 \\ 0 & 0 & 1 \end{vmatrix} \\ &= \begin{vmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix}, \end{aligned}$$

$$\int_0^t S(\tau)\,d\tau = \begin{vmatrix} 0 & -t & 0 \\ t & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix},$$

$$\left(\int_0^t S(\tau)\,d\tau\right)^{2k} = (-1)^k \begin{vmatrix} t^{2k} & 0 & 0 \\ 0 & t^{2k} & 0 \\ 0 & 0 & 0 \end{vmatrix}, \qquad (k \geq 1),$$

$$\left(\int_0^t S(\tau)\,d\tau\right)^{2k+1} = (-1)^k \begin{vmatrix} 0 & -t^{2k+1} & 0 \\ t^{2k+1} & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix}, \qquad (k \geq 0),$$

$$R(t) = \exp\left(\int_0^t S(\tau)\,d\tau\right) = \sum_{k \geq 0} \frac{1}{k!}\left(\int_0^t S(\tau)\,d\tau\right)^k$$

$$= I + \sum_{k \geq 1} \frac{(-1)^k}{(2k)!} \begin{vmatrix} t^{2k} & 0 & 0 \\ 0 & t^{2k} & 0 \\ 0 & 0 & 0 \end{vmatrix} + \sum_{k \geq 0} \frac{(-1)^k}{(2k+1)!} \begin{vmatrix} 0 & -t^{2k+1} & 0 \\ t^{2k+1} & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix}$$

$$= \begin{vmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{vmatrix},$$

$$k(t) = w_1 - S(t)P_1$$

$$= \begin{vmatrix} \sin t + \cos t \\ \sin t - \cos t \\ 0 \end{vmatrix} - \begin{vmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix} \begin{vmatrix} \sin t - \cos t \\ -\sin t - \cos t \\ -1 \end{vmatrix}$$

$$= \begin{vmatrix} \sin t + \cos t \\ \sin t - \cos t \\ 0 \end{vmatrix} - \begin{vmatrix} \sin t + \cos t \\ \sin t - \cos t \\ 0 \end{vmatrix}$$

$$= 0,$$

$$T(t) = 0.$$

Therefore, the required motion is

$$R(t) = \begin{vmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{vmatrix},$$

$$T(t) = 0.$$

It could be directly verified that this motion actually produces the given optical flow and depth information. Interestingly enough, using the discrete iterative scheme (68), the same results come out:

$$R_k = R(t_k) = \begin{vmatrix} \cos t_k & -\sin t_k & 0 \\ \sin t_k & \cos t_k & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

$$T_k = T(t_k) = 0.$$

No discrete errors at all! The reason is simple: $S(t)$ is a constant matrix and hence

$$S(t_k) = \int_0^{t_k} S(\tau)\, d\tau = \sum_{i=1}^{k} S(t_{i-1})\Delta_i = S_k,$$
$$R(t_k) = R_k, T(t_k) = T_k.$$

IV. CONCLUSIONS

It is confirmed that given depth and optic flow of four uncoplanar object points, the rigid body motion can be completely determined. Given only optic flow information without knowing depths, how to obtain the instantaneous rigid body motion is a more difficult and meaningful problem. We already proved: When $K \neq 0$, 8 appropriate points can uniquely determine the instantaneous rotation (or the same, S), the instantaneous translational orientation $K/\|K\|$, and the relative depth $z/\|K\|$. When $K = 0$, 6 appropriate points can uniquely determine the instantaneous rotation (or the same, S). See Zhuang and Haralick.

REFERENCES

D. H. Ballard, and O. A. Kimball, Rigid body motion from depth and optical flow, *Comput. Vision Graphics Image Process.* **22** 1983, 95–115.

X. Zhuang and R. M. Haralick, Rigid Body Motion and the Optic Flow Image, in *The First Conf. on Artificial Intelligence Appl.*, Denver, Dec. 6–7, 1984.