WHAT CAN BE SEEN IN A NOISY OPTICAL FLOW FIELD PROJECTED BY A MOVING PLANAR PATCH IN 3D SPACE?

Soo-Chang Pei* and Lin-Gwo LIOU
Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan, R.O.C.

(Received 13 July 1995; in revised form 15 May 1996; received for publication 24 October 1996)

Abstract—In this paper, we would like to propose a brand new interpretation to the so-called "structure-from-motion" (SFM) problem. The optical flow field projected by a moving rigid planar patch in 3D space is our main consideration. Instead of just obtaining an explicit 3D motion/pose solution like the old approaches did before, we focus our attention on analyzing its error sensitivity, uncertainty, and ambiguity from another point of view. Our new method can handle the above error analysis easily. As known well before, the optical flow field projected by a 3D moving planar patch can be completely expressed by eight coefficients (two for second-order, four for first-order, and two for zeroth-order). Based on these flow coefficients easily determined by a linear regression method or other similar approaches, the error sensitivity of 3D estimates can be analyzed quantitatively and qualitatively in a coarse-to-fine way. The concepts of camera fixation and singular value decomposition (SVD) play important roles in our analysis. There are three goals for our experiments: (1) To prove the correctness of the algorithms (simulated image). (2) To show the tendency of error sensitivity when the 3D poses of the target planar patch are varied in a controlled manner (simulated image). (3) To show that our analysis is workable in the real-world application (real-world image). © 1997 Pattern Recognition Society. Published by Elsevier Science Ltd.

Optical flow field Camera fixation Impacting time Affine transform
Perspective projection

1. INTRODUCTION

The structure-from-motion (SFM) problem has received considerable attention lately. Its purpose is to recover the 3D motion/structure of a moving object from the change of its projected images. Many methods using image points, lines, contours, optical flow, or normal flow have been developed. In this paper, the optical flow field projected by a moving rigid planar patch (MRPP) in 3D space is our main consideration. The optical flow field used here is defined as the instantaneous positional changes of 2D image points, the same as the definition of a so-called "motion flow field." Although there are several slight differences between the traditional definition of an optical flow field and that of a motion flow field, we do not want to emphasize these differences here and confuse readers.

There has been much research[1-22] in estimating the motion/pose of an MRPP in 3D space. For example, using the scaled orthographic projection, there were several area-based (or contour-based) methods: Kanatani[8], Mukundan[9-11], Young[10], and Pei[12,13], approximated the change of projected image shapes by a linear (affine) transform. Under the condition that the MRPP is positioned relatively far enough from the camera, the smaller higher-order change of projected image shapes is completely ignored in their analysis.

However, there were other approaches[1-7] which adopted the perspective projection (pin-hole camera) as their main image model. In 1985, Waxman et al.[1,2] developed the so-called velocity functional method to parameterize the flow field projected by a 3D surface. They concluded that the true optical flow field projected by an MRPP can be completely expressed by eight parameters (two for second-order, four for first-order, and two for zeroth-order). Kanatani[7] directly used the above flow coefficients to solve 3D motion/pose of the planar patch. His analysis was described in terms of complex algebra. Of course there were other methods using different image measurements. For example, Kanatani[3] used line and surface integrals; Tsai and Huang[4-6] used four co-planar point features and solved an SVD problem. The most important conclusion made by them is: there are usually two ambiguous solutions (by at least four point correspondences) when solving the 3D motion/pose of an unknown MRPP.

In fact, ambiguous solutions always happen no matter what kinds of image measurements are used. Some researchers[15-22] are especially interested in the multiple interpretations of an image flow field projected by an MRPP. They also have a similar conclusion: two planar surfaces undergoing different motions may give rise to the same image motion.

However, most of the above researchers[1-22] focused their attention on the complete determination of 3D parameters, while ignoring the analysis of error sensitivity and uncertainty. As known well from experience, the measurement accuracy of the higher-order flow coefficients is seldom reliable enough for an accurate 3D estimate due to the finite image resolution and noise-

* Author to whom correspondence should be addressed.

Pergamon

Pattern Recognition, Vol. 30, No. 9, pp. 1401-1413, 1997
© 1997 Pattern Recognition Society. Published by Elsevier Science Ltd.
Printed in Great Britain. All rights reserved
0031-3203/97 $17.00+0.00
sensitive flow measurement. Old approaches just told us the 3D estimates will be poor when the noise is high. But how poor they are? Can it be described quantitatively? How do the 3D poses of an MRPP affect the accuracy of 3D estimates? Does there exist any singular case when solving the unknowns? What can be seen in a (noisy) optical flow field projected by an MRPP? These questions were rarely answered before!

Therefore, in this paper, a new algorithm is proposed to analyze this sensitivity problem from another point of view. Our main concern is to make a quantitative and qualitative analysis of the error sensitivity, not to develop a more robust or simpler algorithm than the old ones.

The concepts of camera fixation and singular value decomposition (SVD) play important roles in our analysis. First, the process of camera fixation described in Sections 2.1 and 2.2 transforms any given flow field (projected by an MRPP) into its standard form and then reduces the complexity of mathematical derivations. As we will see in Section 2.3, it needs only six coefficients (two for second-order and four for first-order) to represent a standard flow field. Based on the sequential use of the first- and second-order flow coefficients, we will make a "coarse-to-fine" analysis of the problem. In Section 3.1, we will show how the SVD of a $2 \times 2$ matrix $A$ composed of the first-order flow coefficients indicates the possible range of 3D estimates. Then in Section 3.2, we will show how to obtain two deterministic ambiguous solutions when using the second-order flow coefficients. Section 4 makes a global discussion about the uncertainty, ambiguity, and error sensitivity of the 3D estimates. Section 5 designs several sets of experiments to achieve three goals: (1) To justify the correctness of the algorithms (simulated image). (2) To show the tendency of error sensitivity when the 3D poses of the target planar patch are varied in a controlled manner (simulated image). (3) To show that our analysis is workable in the real-world application (real-world image). Finally, Section 6 gives conclusions.

2. STANDARD FORM OF AN OPTICAL FLOW FIELD

2.1. Coordinate transform

Let us see the configuration shown in Fig. 1. An MRPP in 3D space projects an optical flow field on the image plane (see the lower-right corner of Fig. 1). Its plane equation is

$$mxPx + mPy + mzPz = m \cdot P = d.$$  

(1)

Here the symbol $\cdot$ denotes the inner product of vectors. $P = (P_x, P_y, P_z)^T$ is a point on this planar patch. The image point of $P$ is denoted by $p = P/P_z = (P_x, P_y, 1)^T$ (perspective projection). Its optical flow vector is written by $\dot{p} = (p_x, p_y, 0)^T$.

The instantaneous 3D motion of this MRPP can be represented by

$$\dot{P} = \frac{d}{dt} P = \omega \times (P - P_0) + \mathbf{T}.$$  

(2)

The symbol $\times$ denotes the outer product of vectors, $\omega$ and $T$ are instantaneous rotation and translation vectors ($3 \times 1$). $P_0 = (P_{0x}, P_{0y}, P_{0z})^T$ is a specially chosen point on the MRPP, whose image point and optical flow vector are separately denoted by $p_0 = (p_{0x}, p_{0y}, 1)^T$ and $\dot{p}_0 = (p_{0x}, p_{0y}, 0)^T$. Without loss of generality, the global coordinate system is defined to be the same as the camera coordinate system given above.

Now we will introduce the concept of camera fixation. Let us see the configuration shown in Fig. 2 and compare it with that in Fig. 1. The old camera is now rotated (around its own focal point $F$) to a new pose $R$ with an instantaneous rotation velocity $\Omega$ in order that the new image point of $P_0$ is now kept at the image center $(0,0)$ and its new optical flow vector is also kept to zero. Here the $3 \times 3$ pose matrix $R$ is defined as $[U_1 \ U_2 \ U_3]$. $\{U_i\}$s are three $3 \times 1$ orthonormal vectors. The time derivative of $U_i$ is $\dot{U}_i = \Omega \times U_i$. The coordinate transform between the old camera system and the new one is represented by

$$P = RQ.$$  

(3)

A point $P$ seen by the old camera system is now renamed as $Q$ in the new camera system.

The same MRPP in 3D space projects a new optical flow field on the new image plane (see the lower-right
What can be seen in a noisy optical flow field?

Corner of Fig. 2). The plane equation seen by the new camera system is

$$nxQx + nyQy + nzQz \equiv n \cdot Q = d.$$  \hspace{1cm} (4)

Here n = R'Tm. The image point of Q is denoted by q = Q/Qz = (qx, qy, 1)T (perspective projection). Its optical flow vector is written as $\mathbf{q} = (q_x, q_y, 0)^T$. According to the definition of camera fixation, the two constraints, $q_0 = (0, 0, 1)^T$ and $q_0 = (0, 0, 0)^T$, must be satisfied.

Now the instantaneous 3D motion of the MRPP seen by this new camera can be represented by the following special form:

$$\mathbf{Q} = \mathbf{w}^* \times (\mathbf{Q} - \mathbf{Q}_0) + \mathbf{T}_z,$$  \hspace{1cm} (5)

where $\mathbf{z} = (0, 0, 1)^T$; $\mathbf{Q}_0$ is the Z-component of $\mathbf{Q}_0$; $\mathbf{w}^*$ is a new $3 \times 1$ rotation vector; $\mathbf{T}_z$ is a $1 \times 1$ translational component.

In addition, several important transforms are listed below. To improve the reading, we neglect all of their details and show their final results only:

- **Image point.** The transform relationship between p and q is

  $$\mathbf{q} = \begin{bmatrix} q_x \\ q_y \\ 1 \end{bmatrix} = \begin{bmatrix} (u_1 \cdot \mathbf{p})/(u_3 \cdot \mathbf{p}) \\ (u_2 \cdot \mathbf{p})/(u_3 \cdot \mathbf{p}) \\ 1 \end{bmatrix}.$$  \hspace{1cm} (6)

- **Optical flow.** The transform relationship between $\mathbf{q}$ and $\mathbf{q}'$ is

  $$\mathbf{q}' = \begin{bmatrix} q' \\ q'_y \end{bmatrix} = \begin{bmatrix} 1 & 0 & -(u_1 \cdot \mathbf{p})/(u_3 \cdot \mathbf{p}) \\ 0 & 1 & -(u_2 \cdot \mathbf{p})/(u_3 \cdot \mathbf{p}) \\ 0 & 0 & 0 \end{bmatrix} [\mathbf{Q}_0 \times (\mathbf{R} \mathbf{p}) + \mathbf{R} \mathbf{p}].$$  \hspace{1cm} (7)

- **3D motion parameters.** The transform relationship between $\{\mathbf{w}, \mathbf{T}\}$ and $\{\mathbf{w}^*, \mathbf{T}_z\}$ is

  $$\mathbf{w} = \mathbf{R}^* \mathbf{w}^* + \Omega; \quad \mathbf{T} = \mathbf{T}_z^* \mathbf{u}_3 + \mathbf{Q}_0 \Omega (\mathbf{Q} \times \mathbf{u}_3).$$  \hspace{1cm} (8)

From equations (6) and (7), we know that the new image points $\mathbf{q}'$ and their new flow vectors $\mathbf{q}'$ can be directly generated by the old $\mathbf{p}$ and $\mathbf{q}$ once the pose matrix $\mathbf{R}$ and instantaneous rotation $\Omega$ of the new camera are given. We call $\{\mathbf{q}, \mathbf{q}'\}$ the “standard form” of $\{\mathbf{p}, \mathbf{q}\}$.

2.2. Determine the required parameters for camera fixation

In order to transform the old flow field into its standard form, $\mathbf{R}$ and $\Omega$ should be appropriately chosen for satisfying $\mathbf{q}_0 = (0, 0, 1)^T$ and $\mathbf{q}_0 = (0, 0, 0)^T$. We finally obtain the following results:

- **Pose matrix $\mathbf{R}$.** The choice of $\mathbf{R}$ is not unique. We list one solution here:

  $$\mathbf{u}_3 = \mathbf{p}_0/\|\mathbf{p}_0\|,$$
  $$\mathbf{u}_2 = ((\mathbf{p}_0 \times \mathbf{z})/\|\mathbf{p}_0 \times \mathbf{z}\|),$$
  $$\mathbf{u}_1 = \mathbf{u}_2 \times \mathbf{u}_3.$$  \hspace{1cm} (9)

- **Instantaneous rotation $\Omega$.** Substituting the constraints of $\mathbf{q}_0 = (0, 0, 1)^T$ and $\mathbf{q}_0 = (0, 0, 0)^T$ into equation (7), the instantaneous rotation $\Omega$ must satisfy

  $$\Omega = \left( -\mathbf{u}_2 \cdot \frac{\mathbf{p}_0}{\|\mathbf{p}_0\|} \right) \mathbf{u}_1 + \left( \mathbf{u}_1 \cdot \frac{\mathbf{p}_0}{\|\mathbf{p}_0\|} \right) \mathbf{u}_2 + \rho \mathbf{u}_3.$$  \hspace{1cm} (10)

Here $\rho$ can be any real value. We usually set $\rho$ to zero for simplicity.

From equations (9) and (10), $\mathbf{p}_0$ and $\mathbf{p}_0$, we can easily determine the required parameters $\{\mathbf{R}, \Omega\}$ for camera fixation. In fact, mechanically rotating the camera according to the solved $\mathbf{R}$ and $\Omega$ is completely unnecessary! The standard flow field $\{\mathbf{q}, \mathbf{q}'\}$ can be mathematically generated by using equations (6), (7) and (10).

The description of how to transform any optical flow field into its equivalent standard form is now complete.

2.3. Parameterization of a standard flow field

Seen by the fixating camera, the flow field projected by an MRPP in 3D space can be expressed by six coefficients $\{a_i\}$ as follows:

$$\mathbf{q}' = \begin{bmatrix} a_1q_x^2 + a_2q_xq_y + a_3q_x + a_4q_y \\ a_2q_x^2 + a_3q_xq_y + a_4q_x + a_5q_y \end{bmatrix},$$  \hspace{1cm} (11)

where the flow coefficients $a_i$ are defined as

$$a_1 = \omega_x; \quad a_2 = \omega_x; \quad a_3 = \omega_x; \quad a_4 = \omega_x; \quad a_5 = \omega_x; \quad a_6 = \omega_x.$$  \hspace{1cm} (12)

Here $\mathbf{p}$ can be any real value. We usually set $\rho$ to zero for simplicity.

From equations (9) and (10), $\mathbf{p}_0$ and $\mathbf{p}_0$, we can easily determine the required parameters (R and $\Omega$) for camera fixation. In fact, mechanically rotating the camera according to the solved $\mathbf{R}$ and $\Omega$ is completely unnecessary! The standard flow field $\{\mathbf{q}, \mathbf{q}'\}$ can be mathematically generated by using equations (6), (7) and (10).

The description of how to transform any optical flow field into its equivalent standard form is now complete.

3. MAIN ALGORITHM

If $\{\mathbf{w}^*, \mathbf{T}_z^*\}$ can be solved by using the flow coefficients $a_i$, it is easy to transform these 3D motion parameters to the old ones $\{\mathbf{w}, \mathbf{T}\}$ via equation (8). Therefore, without loss of generality, we only analyze the standard form of the flow field described by equation (11).

3.1. Motion estimation from first-order flow coefficients

When measuring the flow coefficients, the higher-order coefficients are always much more error sensitive than the lower-order coefficients. However, a complete solution of 3D unknowns requires the information of these higher-order coefficients. It means that the 3D estimates may not be robust enough. What can be gained if we only use the first-order flow coefficients? This topic will be discussed here.
Let us define a $2 \times 2$ matrix $A$ as
\[
A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} -\omega_{n'y} - n'_{y} - T_{Z} \\ \omega_{n'y} + \omega_{n'y} - T_{Z} \end{bmatrix}.
\] (13)

By singular value decomposition (SVD), the matrix $A$ can be rewritten by
\[
A = USV^{T} = [u_{1}|u_{2}][\lambda_{1} 0 \lambda_{2}][v_{1}|v_{2}]^{T},
\] (14)
where $\lambda_{1} \geq \lambda_{2}$ and $\lambda_{1} \geq 0$. $U$ and $V$ are $2 \times 2$ orthogonal matrices satisfying
\[
u_{2} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}u_{1}; \quad v_{2} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}v_{1}.
\] (15)

From equation (2), $A$ can be further decomposed into the sum of two matrices $A_{1}$ and $A_{2}$, where
\[
A_{1} \equiv \lambda_{2}UV^{T} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix};
\] (16)
\[
A_{2} \equiv (\lambda_{1} - \lambda_{2})u_{1}v_{1}^{T} = \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix}.
\]

It is easy to find that the matrix $A_{1}$ is a scaled orthogonal matrix and matrix $A_{2}$ is not of full rank (rank=1, when $\lambda_{1} \neq \lambda_{2}$).

Besides, from equation (13), $A$ can also be written as the sum of two matrices, $B_{1}$ and $B_{2}$, which are defined as
\[
B_{1} \equiv \begin{bmatrix} -(T_{Z} - k_{1}) & -(\omega'_{y} - k_{2}) \\ -(\omega'_{y} - k_{2}) & -(T_{Z} - k_{1}) \end{bmatrix};
\] (17)
\[
B_{2} \equiv \begin{bmatrix} -\omega_{n'y} - k_{1} & -\omega_{n'y} - k_{2} \\ \omega_{n'y} + k_{2} & \omega_{n'y} - k_{1} \end{bmatrix}.
\]

Here, $k_{1}$ and $k_{2}$ are any two real values. It is obvious that the matrix $B_{1}$ is also a scaled orthogonal matrix like $A_{1}$.

If we want $A_{1} = B_{1}$ and $A_{2} = B_{2}$, what rules should $k_{1}$ and $k_{2}$ obey? It is clear that the determinant of $B_{1}$ must be zero, just like that of $A_{2}$. After some mathematical manipulations, we find that $k_{1}$ and $k_{2}$ must be on a circle characterized by
\[
k_{1} + \frac{\left(\beta_{11} + \beta_{22}\right)}{2} = k_{2} + \frac{\left(\beta_{12} - \beta_{21}\right)}{2}.
\] (18)

Notice that this circle also passes through the origin (0,0). Therefore, $(k_{1},k_{2})$ can be represented by a $\theta$-parameter family: $(k_{1},k_{2})=(k_{1c},k_{2c}) + r\cos \theta, \sin \theta$, where $\theta$ ranges from $0$ to $2\pi$.

From equations (16)-(18), we have
\[
T'_{Z} = -\alpha_{11} + k_{1}, \quad \omega'_{y} = \alpha_{21} + k_{2},
\] (19)
\[
\omega_{n'y} = \beta_{21} - k_{2}, \quad \omega'_{n'y} = \beta_{22} + k_{1},
\]
\[
\omega_{n'y} = -\beta_{11} - k_{1}, \quad \omega'_{n'y} = -\beta_{12} - k_{2}.
\]

We may further separately define two new variables, $h_{n}$ and $h_{n}$, as the amplitudes of $\omega_{n'y}$ ($=[\omega'_{x}', \omega'_{y}]^{T}$) and $n'_{n}$ ($=[n'_{x}, n'_{y}]^{T}$).
\[
\omega_{n'y} \equiv h_{n}\omega_{n'y}; \quad n'_{n} \equiv h_{n}n'_{n},
\] (20)

where $\omega_{n'y}$ and $n'_{n}$ are unit vectors.

After considering equations (18)-(20), it is easy to find that the value $h_{n}h_{n}$ is independent of $k_{1}$ and $k_{2}$:
\[
h \equiv h_{n}h_{n} = \lambda_{1} - \lambda_{2} \geq 0.
\] (21)

Substituting the $\theta$-parameter expression of $(k_{1},k_{2})$ into equation (19), we have
\[
\begin{bmatrix} T'_{Z} \\ \omega'_{y} \end{bmatrix} = \begin{bmatrix} \alpha_{11} - (\beta_{11} + \beta_{22})/2 + r \cos \theta \\ \alpha_{21} - (\beta_{12} - \beta_{21})/2 + r \sin \theta \end{bmatrix}
\] (22)
and
\[
\begin{bmatrix} \omega'_{x} \\ \omega_{n'y} \end{bmatrix} = \begin{bmatrix} h_{n} \cdot \text{sgn}(\cos \theta) \left(\cos \tau_{2} \sin \tau_{1}\right) \\ h_{n} \cdot \text{sgn}(\cos \theta) \left(\cos \tau_{2} \sin \tau_{1}\right) \end{bmatrix},
\] (23)

where $\tau_{1} = [\theta - \varphi + \pi]/2$, $\tau_{2} = [-\theta - \varphi]/2$, $\cos \varphi = (-\beta_{11} + \beta_{22})/(2\pi)$, $\sin \varphi = (\beta_{12} + \beta_{21})/(2\pi)$, $s=\pm 1$, and the $\text{sgn}()$ function is especially defined as
\[
\text{sgn}(x(\theta)) = \begin{cases} 1 & \text{if } x > 0 \\ -1 & \text{if } x < 0 \\ \text{sgn}(x(\theta + \Delta)) & \text{if } x = 0.
\end{cases}
\] (24)

The value of $\Delta$ is a very small positive number.

From equations (21)-(24), we have the following conclusions (only the first-order flow coefficients are used):

- The locus of $(T'_{Z}, \omega'_{y})$ (as a function of $\theta$) forms a circle with center $(T'_{Z}, \omega'_{y})$ and radius $r$. To any assigned value of $\theta$, there is a corresponding unique solution for $(T'_{Z}, \omega'_{y})$. The circle shown in Fig. 3(a) is called the constraint circle for $(T'_{Z}, \omega'_{y})$.
- The locus of the unit vector $\omega_{n'y}$ (as a function of $\theta$) is composed of two opposite half-circles ($s=\pm 1$) which completely occupy the whole unit circle [see Fig. 3(b)]. The unit vector $n'_{n}$ has a similar behavior [see Fig. 3(c)]. For any assigned value of $\theta$, both $\omega_{n'y}$ and $n'_{n}$ are determined to a common sign $s (=\pm 1)$.
- Although $h_{n}$ and $h_{n}$ have not yet been completely solved here, we know the value of $h_{n}$ which is defined as $h_{n}h_{n}$. Equation (21) indicates that $h_{n}h_{n}$ can be easily obtained once giving $h_{n}h_{n}$. In fact, $h_{n}h_{n}$ (or $h_{n}$) is usually of finite value. For example, it is impossible for an MRPP to move in 3D space by an infinitely large 3D rotation. Besides, the case that $h_{n}h_{n}$ is not acceptable because the projected image of the MRPP is now reduced to a line segment. Both $h_{n}$ and $h_{n}$ should lie in some finite range [see Fig. 3(d)]. Therefore, as illustrated in Fig. 3(a)-(d), the possible 3D locus of $\omega'_{y}$ (or $n'$) must also occupy a limited region in...
What can be seen in a noisy optical flow field?

3D space. Figure 3(e) shows a hollow cylinder which contains all of the possible values of $\omega^*$ (a twisted, closed 3D slice). The height $r$ and radius $h_n$ of this cylinder are controlled by the difference of the two singular values, $\lambda_1$ and $\lambda_2$. In a similar way, Fig. 3(f) shows a circular band on a unit hemisphere, which indicates all of the possible directions $n$ of the planar patch. From Fig. 3(a)–(f), readers can see how the uncertainty of 3D estimates can be expressed quantitatively when using only first-order coefficients.

- Without the information of the second-order flow coefficients, the whole set of 3D motion/pose estimates are in fact a family of two unknown variables, $\theta$ and $h_n$. The inherent depth indeterminacy $Q_{\omega}$ in a monocular vision system is not considered here. Therefore, from the above results, only partial 3D
3.2. Motion estimation from second-order flow coefficients

Section 3.1 concludes that the 3D motion/pose of an MRPP in 3D space cannot be uniquely solved by its first-order flow coefficients only. If its second-order flow coefficients are used, can we obtain an unique solution? This topic will be discussed here.

From the definitions of \(a_1\) and \(a_2\) in equation (12), we have

\[
\begin{bmatrix}
\omega_y^* - \hat{\rho}_x^* T_z^* \\
\omega_z^* - \hat{\rho}_x^* T_z^*
\end{bmatrix}
\begin{bmatrix}
h_y \\
h_z
\end{bmatrix} =
\begin{bmatrix}
a_1 \\
a_2
\end{bmatrix}
\tag{25}
\]

and

\[
\begin{bmatrix}
h_z \\
h_n
\end{bmatrix} =
\frac{1}{-T_z^* (\omega_y^* n_y^* + \omega_z^* n_z^*)}
\begin{bmatrix}
T_z^* (-a_1 \hat{\rho}_x^* + a_2 \hat{\rho}_y^*) \\
a_2 \omega_y^* + a_1 \omega_z^*
\end{bmatrix}.
\tag{26}
\]

Because the constraint of equation (21), that is \(h_n = h\), must be satisfied, we may substitute equations (26) and (19) and \((k_1, k_2) = (k_{11}, k_2) + r (\cos \theta, \sin \theta)\) into equation (21). It finally becomes a third-order polynomial of \(z = \cos \theta\)

\[
f(z) = z^3 + \left(-e_1 \right) z^2 + \left(-1 - \frac{a_1^2 + a_2^2}{4r^2}\right) z + e_2 = 0,
\tag{27}
\]

where

\[
e_1 = \frac{\beta_1 + \beta_2}{2} + \alpha_1;
\]

\[
e_2 = \left(\frac{a_2^2 - a_1^2}{4r^2}\right) \cos \varphi + \left(\frac{a_1 a_2}{2r^2}\right) \sin \varphi + \left(\frac{e_1}{r}\right).
\tag{28}
\]

Let us see Fig. 4(a). Although the polynomial \(f(z) = 0\) has theoretically three roots for \(z\), there is usually only one root, say \(z = z_0 = \cos \theta_0\), which can satisfy \(-1 \leq z \leq 1\). To a value \(z_0\), the angle \(\theta\) described in equation (22) can be either \(\theta_0\) or \(-\theta_0\). Let us see Fig. 4(b). Solving \(f(z) = 0\) is equivalent to drawing a corresponding vertical line on the \((T_z^*, \omega_z^*)\) plane and then finding its two intersecting points on the constraint circle.

While using equation (27), the case that \(r = h = 0\) will make equation (27) unusable and we have to consider equation (25) again. Even so, we still have two solutions:

1. If \(h_n = 0\) (that is, \(\omega_{xy}^* = [0, 0]^T\)), the vector \(n_{xy}^*\) is equal to \(-[a_1, a_2]^T/T_z^*\).
2. If \(h_n = 0\) (or \(n_{xy}^* = [0, 0]^T\)), the vector \(\omega_{xy}^*\) is equal to \([-a_2, a_1]^T\).

4. DISCUSSIONS

In this section, we would like to make a global description about what can be seen in a noisy flow field projected by an MRPP. It is hoped that the readers can get a clearer picture to this problem.

4.1. Parameterization of noisy flow measurement

Our analysis is totally based on the coefficients of the (standard) flow field. These coefficients can be determined from various sources: (1) the optical flow estimated by the so-called gradient-based methods, (2) the normal flow estimated by moving image contours, (3) the displacement vectors estimated by the so-called block-matching or correlation-based methods; the discrete time image measurement is now an approximation of the continuous one. No matter where these coefficients come from, our analysis can be directly applied to them.

The most important reason for us to adopt flow coefficients as basic measures is the requirement for reliability. Due to the redundancy of input data, the accuracy of flow coefficients is much better than that of individual image points and optical flow vectors. So
What can be seen in a noisy optical flow field?

1. The uncertainty region of a noisy flow field (for its 3D estimates) is determined by the intersection of two areas: (1) the possible perturbation range of the vertical line, (2) the possible perturbation range of the constraint circle.

2. From first-order coefficients:
   - Vertical line from second-order coefficients:
   - Constraint circle from first-order coefficients:

3. The accuracy of 3D estimates will be improved without any question.

4.2. Uncertainty of 3D estimates

From Section 3.1 and Fig. 3, we know that the SVD of A dominates our interpretation to 3D motion/pose. The most important quantity is the difference between singular values, \( \lambda_1 - \lambda_2 \). It specifies the size of uncertainty when we only consider the first-order flow coefficients.

From Section 3.2 and Fig. 4, it is straightforward that solving the SFM problem is just equivalent to drawing a vertical line through the constraint circle of \((T_z', \omega_z')\) to find its two intersecting points. The constraint circle is estimated by the four first-order coefficients only, but the vertical line is estimated by the second-order ones. Because the second-order flow coefficients are usually much more error-sensitive than the first-order coefficients, the positional perturbation of the vertical line (due to noise) is often larger than that of the constraint circle. Figure 5 illustrates the above idea. So the final 3D estimates from a noisy flow field must fall into an uncertainty region specified in Fig. 5, whose size depends on the amplitude of noise.

On the other hand, a quantity called impacting time, denoted by \( t_{imp} \), is just equal to \( 1/T_2' \) defined in this paper. Readers should notice that both of the two ambiguous solutions have the same value of \( t_{imp} \). It means that we will always have a unique estimate for \( t_{imp} \), no matter how many ambiguous solutions are obtained. The size and position of the uncertainty region shown in Fig. 5 is directly related to the maximum and minimum range of \( t_{imp} \), which can be used as the basis of obstacle avoidance in an auto-vehicle’s vision system.

Some readers may think that a smaller constraint circle for \((T_z', \omega_z')\) often implies a smaller uncertainty for 3D estimates. But it is wrong! Let us see Fig. 6 for further explanation. The same perturbation of the vertical line induces different results: smaller radius \( r \) means smaller variance for \( T_z' \) and \( \omega_z' \), but larger variance for the directions of \( \omega_{xy} \) and \( n_{xy} \). It is because the uncertainty regions for both cases occupy different ranges of \( \theta \) which are directly related to the uncertainty of \( \omega_{xy} \) and \( n_{xy} \) according to equation (23). It seems that there is a "trade-off" between the accuracy of different 3D estimates.

5. EXPERIMENTS

In this section, several sets of experiments are designed for testing our algorithm. They have three goals:

1. To prove the correctness of the algorithms (simulated image).
2. To show the tendency of error sensitivity when the relative poses of the target planar patch are varied in a controlled manner (simulated image).
3. Our analysis is workable in the real-world application (real-world image).

Before describing these experiments, there are two things to be noticed:

- All the planar patches tested in Sections 5.1 and 5.2 are the same. It is a 3 x 3 square patch with 25 uniformly-distributed feature points. Its central point, the 13th feature point, is the specially chosen fixing point \( P_0 \). Parameterization of the given flow field is totally based on these 25 feature points and their optical flow vectors. In order to verify the correctness of our algorithm, the flow vectors in computer simulations are generated by a simple, well-known physical rule:

\[
\begin{bmatrix}
\hat{p}_x \\
\hat{p}_y
\end{bmatrix} = \frac{1}{P_z} \begin{bmatrix}
1 & 0 & -p_z \\
0 & 1 & -p_y
\end{bmatrix} \hat{P}.
\]

Here \( \hat{P} \) is generated by the definition given in equation (2).

- In Section 5.2, the noisy flow field is generated by purposely adding a 2D random noise \((n_1, n_2)\) on the error-free flow field like this

\[
\begin{bmatrix}
\hat{p}_x \\
\hat{p}_y
\end{bmatrix} = \begin{bmatrix}
\hat{p}_x \\
\hat{p}_y
\end{bmatrix} + \begin{bmatrix}
n_1 \\
n_2
\end{bmatrix}.
\]

Here \( n_1 \) and \( n_2 \) are random variables of normal distribution \( N(0, \sigma) \) and

\[
\sigma = \sqrt{\frac{n_1^2 + n_2^2}{2}}.
\]
The percentage ratio $\delta$ is defined as the noise level. For simplicity, other possible error sources such as quantization errors, camera calibration errors, and fixation errors are assumed to be error-free throughout the experiments in Section 5.2.

5.1. Error-free experiments

Figure 7(a) shows an optical flow field $\mathbf{p}$ projected by an MRPP in 3D space. The true 3D motion/pose of the MRPP is randomly assigned. By equations (9) and

![Figure 7](image-url)

Fig. 7. Several plots of the results obtained in the error-free experiments.
What can be seen in a noisy optical flow field?

| true para. | $\omega^* = (-0.2349, -0.1054, 0.4214)$ | $T^* = (0.0000, 0.0000, 1.5613)$ |
| $n = (-0.4069, -0.4939, 0.7684)$ | $Q_0 = (0.0000, 0.0000, 10.4403)$ |
| $a_1 = -0.0262$ | $a_2 = 0.3310$ | $a_3 = -0.2053$ | $a_4 = -0.4892$ |
| $a_5 = 0.5458$ | $a_6 = 0.0014$ | $T_z = 0.1495$ |

using the first-order flow coef. only

$r = 0.1072$ | $h = 0.2144$ | $\varphi = 0.2674$ (rad) |

$T^* = (0.0000, 0.0000, 1.5613)$ | $Q_0 = (0.0000, 0.0000, 10.4403)$ |

using both of the first- and second-order flow coef.

3-order polynomial: $z^3 + (0.9511)z^2 - (3.3985)z + 1.2377 = 0$

three roots: $-2.5047, 0.4439, 1.1096$

| #1 solution | ($T^*_z, \omega^*_z$) = (0.1495, 0.4214) |
| $(\omega^*_x, \omega^*_y) = (-0.2349, -0.1054)$ |
| $n = (-0.4069, -0.4939, 0.7684)$ |

| #2 solution | ($T^*_z, \omega^*_z$) = (0.1495, 0.6135) |
| $(\omega^*_x, \omega^*_y) = (-0.0961, 0.0792)$ |
| $n = (0.3539, -0.7889, 0.5123)$ |

Fig. 8. The results of the error-free experiment.

(10), this flow field is then changed into its standard form shown in Fig. 7(b). Notice that the flow fields shown in Fig. 7(a) and (b) have been suitably scaled (0.3 and 0.5) for a better look. The first part of the table in Fig. 8 lists: (1) the true 3D motion/pose of the MRPP seen by the rotated camera, (2) the true flow coefficients of the standard flow field shown in Fig. 7(b).

If we only consider the first-order flow coefficients, only partial information can be recovered. The second part of the table in Fig. 8 lists all of the solved quantities. Figure 7(c) plots the solved constraint circle. It is easy to check that the true $(T^*_z, \omega^*_z)$ (denoted by "*")) is lies exactly on this circle.

Figure 7(d) and (e) show the possible values of $\omega_x$ and $n_x$. Two opposite half-circles (solid for $s=+1$ and dotted for $s=-1$) occupy the whole unit circle. If the true angle is given and substituted into equation (23), we will obtain its corresponding $\omega_x$ and $n_x$ (represented by a dotted straight line through the origin). It is easy to check that the true solutions (marked by "*") are just lying on the straight lines, which proves equation (23). Figure 7(f) shows the uncertainty of $h_x$ and $h_n$. The true solution (marked by "*")) is just lying on this curve, which proves equation (21).

If the second-order flow coefficients are considered, we will have two ambiguous solutions. Figure 7(g) shows the third-order polynomial function, $f(z) = 0$. It has three roots: $z = \cos \theta = [-2.5047, 0.4439, 1.1096]$. Only the second one satisfies the constraint $|z| \leq 1$. So we have two ambiguous solutions listed in the third part of the table shown in Fig. 8. Notice that the first solution is just equal to the true solution. It proves our algorithm.

5.2. Analysis of error sensitivity

To analyze the error sensitivity of the 3D estimates, five cases are defined here. All of the MRPPs tested in these cases have the same point $P_0$ located at $(0, 0, 10)^T$ and the same 3D motion: $\omega = (0, 0.5, 0.1)^T$, translation $T = (0, 0, 2)^T$. The normal vectors of these five MRPPs can be represented by the following form: $m = (\sin \alpha, 0, \cos \alpha)^T$. From cases 1-5, their corresponding angle $\alpha$ are separately set to $0^\circ$, $15^\circ$, $30^\circ$, $45^\circ$, and $60^\circ$.

For the above five specially-designed testing cases, all of their generated flow fields have been standardized and hence we can directly ignore the process of camera fixation. So $\omega = \omega^*$, $T = T^*$, $m = n$, and $T^* = 2/10 = 0.2$.

Figure 9(a)-(e) show the error performance of the five cases for noise level $\delta$ ranging from 0% to 20%. Each point on the curves is the average of 100 tests. Figure 9(a) shows the position error, $||\Delta c||/||c||$, of the center $c$ of the constraint circle. Figure 9(b) shows the deviations $\Delta r$ of its radius $r$. The radii $r$ of the five cases are: 0.0551, 0.1443, 0.2500, and 0.4330. We find that the estimate of the constraint circle is very robust. Besides, it seems that the different orientations of a planar patch do not affect the error sensitivity of the uncertainty estimate remarkably.

Figure 9(c)-(e) show the results when the second-order flow coefficients are considered. Because there will be two ambiguous solutions, we choose the one nearest to the true solution as our estimate. Figure 9(c) shows the estimation errors, defined as $||\Delta \omega^*||/||\omega^*||$, of
the rotation $\omega^*$. Figure 9(d) shows the estimation errors, defined as $\|\Delta T_x\|/\|T_x^0\|$, of the translation $T = (0, 0, T_z)$. Figure 9(e) shows the estimation errors, $\Delta n$, of the normal vector $n$. We find that the estimates obtained from a more-slanting planar patch are more error sensitive than those from a less-slanting one. This tendency is especially remarkable in Fig. 9(d) when estimating the translation. No doubt, the error performance of the estimate of $T_x^0$ will affect the error performances of the other estimates such as $\omega^*$ and $n$. It is because $\omega^*$ and $n$ are determined after we know the value of $z$ (or $T_z^0$). Although there is a tendency that a larger radius $r$ (of the constraint circle) usually induces a smaller error sensitivity in estimating the directions of $\bar{\omega}^*$ and $\bar{n}^*$, the tendency shown in Fig. 9(d) seems much stronger and then dominates the later error performances shown in Fig. 9(c) and (e). Notice that the differences among the curves in Fig. 9(c) and (e) are not so obvious as that in Fig. 9(d).

5.3. Real-image experiment

To test the performance of our algorithm in a real application, we designed a simple and well-controlled experiment in our laboratory. A planar patch mounted on a pyramid-type object is considered [see Fig. 10(a)]. Each one of the four sides of the pyramid-type object is an equilateral triangle. For convenience in 3D measurement, the relative 3D motion between the camera and the plate is controlled by two parameters: a $Z$-direction translation $v_Z$ ($= -27.5$ cm/s) of the camera, and a rota-
What can be seen in a noisy optical flow field?

Fig. 10. (a) The testing 3D configuration for the real-world experiment. (b) Observed image frame at t=0. The white dots denote the selected feature points on the plate. (c) The point trajectory of every selected feature point in the image sequence (of 7 frames). (d) Estimated flow vector at t=0. Scale of the flow vector is set to 0.5.

The CCD camera has been calibrated by a simple pin-hole model. The resolution of the image plane is $512 \times 480$ (pixels); viewing angle is about $30^\circ$; the focal length is about $18.3$ cm.

The three-dimensional rotation $\theta (=+60^\circ/s)$ around the $Y$-axis which passes through a chosen fixed point $P_0$ on the plate. At $t=0$, the 3D coordinate of $P_0$ is $(0, 0, 109.8$ cm); the plane equation of the plate is $(1/3)Y + (2\sqrt{2}/3)Z = (2\sqrt{2}/3) \cdot 109.8$ cm.
In order to estimate the instantaneous flow vectors at \( t=0 \), an image sequence of seven image frames \((t:-0.5, -0.33, -0.167, 0, +0.167, +0.33, +0.5 \text{ s})\) is taken. Figure 10(b) shows the observed image at \( t=0 \); 23 white dots denote the selected feature points on the plate. After finding the corresponding feature points at each time instant, the trajectory \((x(t),y(t))\) of every feature point \((x(0),y(0))\) is fitted by a third-order polynomial model:

\[
(x(t), y(t)) = C_0 + C_1 t + C_2 t^2 + C_3 t^3
\]

Therefore, the flow vector \((x(0),y(0))\) is estimated by \((C_1,C_2, C_3)\). Figure 10(c) shows the point trajectory; Fig. 10(d) shows the estimated flow vectors. Notice that the length unit of image coordinates in Fig. 10(c) and (d) has been changed from one pixel to one focal length.

Figure 11 lists the final 3D estimates. The #1 solution is quite close to our true 3D measurements. Large magnitude of estimation errors are reasonable because of the errors from (1) camera calibration, (2) image quantization. Besides, the small viewing angle of the object does increase the error sensitivity.

### 6. CONCLUSION

In this paper, we propose a new algorithm to solve and analyze the SFM problem from the optical flow field projected by an MRPP in 3D space. Although our approach has exactly the same solution just like the old approaches did before, there are several additional advantages of our algorithm: (1) It is very easy to analyze the error sensitivity, ambiguity, and uncertainty in a noisy optical flow field. (2) Our algorithm introduces the concept of camera fixation and reduces the number of parameters while representing a flow field (from eight to two ambiguous solutions). (3) The SFM problem can be solved by “levels.” According to the accuracy of observed flow field, we can interpret the flow field from coarse to fine (from constraint circle to two ambiguous solutions). (4) Our derivations clearly show the trade-off of accuracy when solving the 3D motion/pose parameters.

To analyze the SFM problem more completely, several experiments are designed. We draw the following conclusions: (1) When using first-order flow coefficients only, the estimate for the constraint circle is quite error insensitive. (2) When using full information of flow coefficients, the estimate is usually error-sensitive. (3) The tendency of the error performance with respect to different orientations is shown.

What can be seen in a noisy flow field projected by an MRPP in 3D space? Finally we can say: “Much more than just obtaining two ambiguous solutions.”

### REFERENCES

What can be seen in a noisy optical flow field?


About the Author — SOO-CHANG PEI was born in Soo-Auo, Taiwan, R.O.C., on 20 February 1949. He received the B.S. degree from National Taiwan University in 1970 and the M.S. and Ph.D. degrees from the University of California, Santa Barbara, in 1972 and 1975, respectively, all in Electrical Engineering. He was an Engineering Officer in the Chinese Navy Shipyard at Peng Fu Island from 1970 to 1971 and a Research Assistant at the University of California, Santa Barbara, from 1971 to 1975. He was Professor and Chairman in the Department of Electrical Engineering at Tatung Institute of Technology from 1981 to 1983. Presently he is the Professor and Chairman of the Department of Electrical Engineering at National Taiwan University. His research interests include digital signal processing, digital picture processing, optical information processing, lasers, and holography. Dr Pei is a member of the IEEE, Eta Kappa Nu, and the Optical Society of America.

About the Author — LIN-GWO LIOU was born in Taiwan. He received the B.S. degree from National Chiao Tung University (N.C.T.U.) in Taiwan in 1989 and the Ph.D. degree from the National Taiwan University in 1995, both in Electrical Engineering. He is currently doing military service. His research interests include motion image analysis, methods for 3-D object reconstruction, and pattern recognition in image applications.