International Journal of Computer Vision, 11:3, 267–282 (1993)

3-D Interpretation of Optical Flow by Renormalization

KENICHI KANATANI
Department of Computer Science, Gunma University, Kiryu, Gunma 376, Japan

Abstract
This article studies 3-D interpretation of optical flow induced by a general camera motion relative to a surface of general shape. First, we describe, using the "image sphere representation," an analytical procedure that yields an exact solution when the data are exact: we solve the *epipolar equation* written in terms of the *essential parameters* and the *twisted optical flow*. Introducing a simple model of noise, we then show that the solution is "statistically biased." In order to remove the statistical bias, we propose an algorithm called *renormalization*, which automatically adjusts to unknown image noise. A brief discussion is also given to the *critical surface* that yields ambiguous 3-D interpretations and the use of the *image plane representation*.

1 Introduction

This presentation studies computational procedures for computing the 3-D interpretation of optical flow. Since the displacement of each image point is small (theoretically infinitesimal), the computation is unstable and sensitivity affected by image noise if the computation is based on individual displacements. However, the use of optical flow has the advantage that the flow can be detected densely (usually at each pixel) over the entire image. Hence, it is expected that a reliable 3-D interpretation can be obtained by optimization over the entire image frame.

If the surface is expressed as a parameterized equation, say a polynomial or a collection of planar patches, the problem reduces to estimating the coefficients [1, 24, 30]. The problem becomes very difficult if no surface model is assumed. An analytical solution can be obtained if spatial derivatives of the flow velocity are available [15]. An important clue is obtained from the fact that the flow due to camera rotation is depth independent [9], thereby globally continuous and smooth. Hence, a sudden change of the flow over a small number of pixels reflects a depth discontinuity [15, 23]—the change indicates the translational component of the flow (*motion parallax*). More systematically, the translation velocity can be determined by optimally searching for a rotation velocity such that subtraction of the effect of rotation results in a flow that has a common *focus of expansion* [22]. A more direct approach is to do numerical search to minimize the sum of the squares of the differences between the observed flow and the expected theoretical expression [4, 2, 5, 29].

On the other hand, analytical (*linearized*) procedures were found for finite motion using point correspondences between two views [16, 28, 31]. Since optical flow is simply an infinitesimal limit of a finite image motion, an approximation for a small motion yields an algorithm for optical flow, and a linearized procedure for optical flow was proposed by Zhuang et al. [32]. Here, modifying the linearization of Zhuang et al. [32], we transform the algorithm of Weng et al. [31]—who solved the epipolar equation by least squares over a large number of feature points—into an algorithm for an optical flow. Adopting the "image sphere representation," we show that all we need to solve is the *epipolar equation* written in terms of the *essential parameters* and the *twisted optical flow*.

It has been pointed out that the solution based on least-squares minimization is likely to be systematically biased [3]. In this article, we analyze this statistical bias by introducing a simple model of noise; we then present an algorithm called "renormalization," which removes the statistical bias by automatically adjusting to unknown image noise. This is a generalization of the approach of Tagawa et al. [25, 26, 27], who also proposed iterative methods to remove statistical bias. A random-number simulation is given to observe its effectiveness.

We also analyze the numerical instability that occurs when the field of view is very small, and show how the condition number grows as the field of view decreases. Finally, we give a brief discussion of the *critical surface* that yields ambiguous 3-D interpretations in relation to the decomposability of the essential parameters and the use of the *image plane representation.*

2 Optical Flow Equation

Assume the following camera imaging model. The camera is associated with an *XYZ* coordinate system with origin *O* at the center of the lens and *Z*-axis along the optical axis (figure 1a). The plane $Z = f$ is identified with the image plane. We call the origin *O* the *viewpoint*; the constant *f* is often called the *focal length.*

A point on the image plane is represented by the unit vector **m** starting from the viewpoint *O* and pointing toward that point (figure 1b). We call **m** the *N-vector* of the point [10]. The use of N-vectors is equivalent to the use of *homogeneous coordinates*. Mathematically, homogeneous coordinates can be multiplied by any nonzero constant. From a computational point of view, however, it is more convenient to normalize the three component into a unit vector. This is also equivalent to considering a hypothetical "spherical imaging surface" of unit radius centered at the viewpoint *O.* This representation has been (implicitly or explicitly) used by many authors (e.g., [17]) and was fully developed into a mathematical framework by Kanatani [10], who called it *computational projective geometry.*

Consider a surface of general shape. Let $r(\mathbf{m})$ be the distance of the surface point of N-vector **m** from the viewpoint *O.* If the camera moves with rotation velocity ω around the viewpoint and translation velocity ν relative the surface, an optical flow is induced on the image plane (figure 1b). We call the pair $\{\omega, \nu\}$ the *motion parameters*. We also abbreviate the camera motion with motion parameters $\{\omega, \nu\}$ simply as "motion $\{\omega, \nu\}$."

We represent the flow by the time derivative $\dot{\mathbf{m}}$ of the N-vector representing that point (i.e., the velocity on the hypothetical image sphere of unit radius) and call it the *N-velocity.* Since the N-vector **m** is normalized to a unit vector, the N-velocity $\dot{\mathbf{m}}$ is always orthogonal to **m**:

$$(\mathbf{m}, \dot{\mathbf{m}}) = 0 \tag{1}$$

In this article, we use $(\cdot, \cdot)$ to denote inner product of vectors.

The orthogonal projection of a vector **a** onto the plane of unit surface normal **m** is given by $\mathbf{P}_\mathrm{m}\mathbf{a}$ (figure 1b), where

$$\mathbf{P}_\mathrm{m} = \mathbf{I} - \mathbf{m}\mathbf{m}^T \tag{2}$$

This is a *projection matrix*:

$$\mathbf{P}_\mathrm{m}^T = \mathbf{P}_\mathrm{m}, \qquad \mathbf{P}_\mathrm{m}^2 = \mathbf{P}_\mathrm{m} \tag{3}$$

Theorem 1: *Camera motion* $\{\omega, \nu\}$ *relative to a surface of depth* $r(\mathbf{m})$ *induces an optical flow*

$$\dot{\mathbf{m}} = -\omega \times \mathbf{m} - \frac{\mathbf{P}_\mathrm{m}\nu}{r(\mathbf{m})} \tag{4}$$

Proof: Rotation of the camera with rotation velocity ω is equivalent to rotation of the scene with rotation velocity $-\omega$, which causes optical flow $\dot{\mathbf{m}} = -\omega \times \mathbf{m}$. Translation of the camera with velocity ν is equivalent to translation of the scene with velocity $-\nu$.

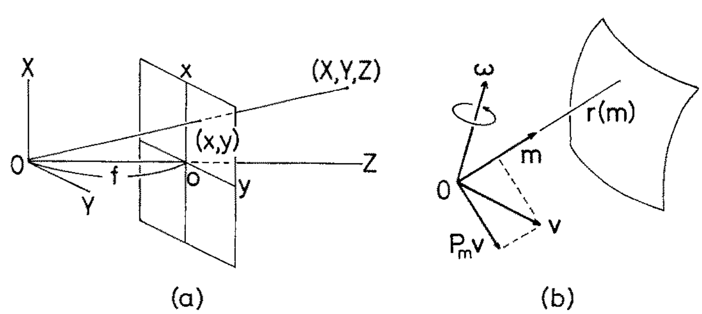

Fig. 1. (a) Camera imaging geometry and N-vector **m**. (b) Depth map $r(\mathbf{m})$ and camera motion $\{\omega, \nu\}$.

Motions along the line of sight do not cause any image motion, so only the orthogonal projection $-\mathbf{P}_m\boldsymbol{\nu}$ is perceived (figure 1b) and is inversely scaled by the distance $r(\mathbf{m})$. Hence, $\dot{\mathbf{m}} = -\mathbf{P}_m\boldsymbol{\nu}/r(\mathbf{m})$. The total flow is the sum of these two flow components. □

Equation (4) was first used by Maybank [17] to analyze the uniqueness of 3-D interpretation, and is mathematically equivalent to the image-coordinate form derived by many researchers.

3 3-D Interpretation of Optical Flow

We assume that a flow field $\dot{\mathbf{m}}(\mathbf{m})$ is defined over a solid angle $\Omega(S)$ corresponding to the field of view S (figure 1b). From equation (4), we observe that if $\{\omega, \nu\}$ and $r(\mathbf{m})$ are a solution, so are $\{\omega, k\nu\}$ and $kr(\mathbf{m})$ for an arbitrary nonzero constant k, meaning that a large motion far from the viewer is indistinguishable from a small motion near the viewer. This is inevitable because the N-velocity $\dot{\mathbf{m}}$ (= the rate of change of the line of sight) is nondimensional: we cannot derive quantities involving length from nondimensional data. Since it is easy to check if $\nu = \mathbf{0}$ (appendix A), we assume that $\nu \neq \mathbf{0}$ has already been confirmed, and normalize ν into a unit vector. The sign of ν is still indeterminate: if $\{\omega, \nu\}$ and $r(\mathbf{m})$ are a solution, so are $\{\omega, -\nu\}$ and $-r(\mathbf{m})$. However, one predicts a positive depth and the other a negative depth, so we can pick out the correct sign.

If the motion parameters $\{\omega, \nu\}$ are correctly estimated, the depth $r(\mathbf{m})$ is easily determined [4, 2, 17]. Consider the least-squares criterion

$$\left\| \dot{\mathbf{m}} + \omega \times \mathbf{m} + \frac{\mathbf{P}_m\nu}{r(\mathbf{m})} \right\|^2 \rightarrow \min \quad (5)$$

Differentiating this with respect to $1/r(\mathbf{m})$, setting the result to 0, and solving it, we obtain

$$r(\mathbf{m}) = \frac{1 - (\mathbf{m}, \nu)^2}{|\mathbf{m}, \omega, \nu| - (\dot{\mathbf{m}}, \nu)} \quad (6)$$

where $|\mathbf{a}, \mathbf{b}, \mathbf{c}|$ $\big(= (\mathbf{a}, \mathbf{b} \times \mathbf{c}) = (\mathbf{b}, \mathbf{c} \times \mathbf{a}) = (\mathbf{c}, \mathbf{a} \times \mathbf{b})\big)$ denotes the scalar triple product of vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$. Let us call equation (6) the *motion parallax equation*. From this, the negative depth solution can be excluded. If $r(\mathbf{m}) > 0$ cannot be imposed on every $\mathbf{m} \in \Omega(S)$ in the presence of noise, it is reasonable to ask for the "majority vote":

$$\int_{\Omega(S)} [|\mathbf{m}, \omega, \nu| - (\dot{\mathbf{m}}, \nu)]\, d\Omega(\mathbf{m}) > 0 \quad (7)$$

Here, $\int_{\Omega(S)} d\Omega(\mathbf{m})$ denotes integration over the solid angle $\Omega(S)$, but this is a mere notational convenience—it is understood to mean the summation over all the pixels at which the flow is defined. In real applications, the summation should be weighted by some certainty measure that assigns the reliability of the detected flow, since flow detection is reliable in highly textured regions but unreliable in regions with almost homogeneous gray levels.

4 Twisted Flow and the Epipolar Equation

Instead of analyzing the flow field $\dot{\mathbf{m}}(\mathbf{m})$, we rotate the N-velocity $\dot{\mathbf{m}}$ by 90° around each $\mathbf{m}$ (figure 2) and call the resulting flow field $\dot{\mathbf{m}}^*(\mathbf{m})$ the *twisted flow*. Since the N-velocity $\dot{\mathbf{m}}$ is orthogonal to $\mathbf{m}$ (equation (1)), the twisted flow $\dot{\mathbf{m}}^*$ is given by

$$\dot{\mathbf{m}}^* = \mathbf{m} \times \dot{\mathbf{m}} \quad (8)$$

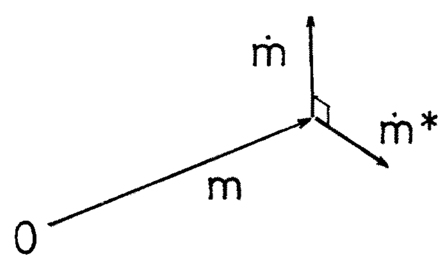

Fig. 2. Twisted flow $\dot{\mathbf{m}}^*$.

If we define a symmetric matrix

$$\mathbf{K} = (\omega, \nu)\,\mathbf{I} - \frac{1}{2}(\omega\nu^\top + \nu\omega^\top) \quad (9)$$

the following relationship holds:

$$(\mathbf{m}, \mathbf{K}\mathbf{m}) = (\omega, \mathbf{P}_m\nu) = (\nu, \mathbf{P}_m\omega) \quad (10)$$

We call the pair $\{\mathbf{K}, \nu\}$ the *essential parameters*. We say that a pair $\{\mathbf{K}, \nu\}$ consisting of a symmetric matrix and a unit vector ν is *decomposable* if there exists a vector ω such that $\mathbf{K}$ has the form of equation (9).

Proposition 1. *A pair $\{\mathbf{K}, \nu\}$ is decomposable if and only if*

$$(\nu, \mathbf{K}\nu) = 0$$

$$\mathbf{K} = \frac{1}{2}(\text{tr } \mathbf{K})(\mathbf{I} - \nu\nu^T) + \mathbf{K}\nu\nu^T + \nu\nu^T \mathbf{K} \tag{11}$$

Proof: If $\mathbf{K}$ has the form of equation (9) for some vector ω, it is easy to confirm equations (11) by direct substitution. Conversely, if $\mathbf{K}$ satisfies equations (11), define

$$\omega = \frac{1}{2}(\text{tr } \mathbf{K})\nu - 2K\nu \tag{12}$$

Then, it is easy to confirm that $\mathbf{K}$ is expressed in the form of equation (9). □

Proposition 2. *The twisted flow* $\dot{\mathbf{m}}^*$ *induced by camera motion* $\{\omega, \nu\}$ *relative to a surface of depth* $r(\mathbf{m})$ *has the form*

$$\dot{\mathbf{m}}^* = -\mathbf{P}_m\omega - \frac{\mathbf{m} \times \nu}{r(\mathbf{m})} \tag{13}$$

Proof: From theorem 1 and equation (8), we have

$$\dot{\mathbf{m}}^* = -\mathbf{m} \times (\omega \times \mathbf{m}) - \frac{\mathbf{M} \times \mathbf{P}_m\nu}{r(\mathbf{m})} \tag{14}$$

It is easy to confirm that $\mathbf{m} \times (\omega \times \mathbf{m}) = \mathbf{P}_m\omega$ and $\mathbf{m} \times \mathbf{P}_m\nu = \mathbf{m} \times \nu$. □

Theorem 2: *A flow field* $\dot{\mathbf{m}}$ *defines a depth* $r(\mathbf{m})$ *if and only if there exist essential parameters* $\{\mathbf{K}, \nu\}$ *such that*

$$(\dot{\mathbf{m}}^*, \nu) + (\mathbf{m}, \mathbf{K}\mathbf{m}) = 0, \quad \mathbf{m} \in \Omega(S) \tag{15}$$

Proof: Equation (13) is written as

$$\dot{\mathbf{m}}^* + \mathbf{P}_m\omega = -\frac{\mathbf{m} \times \nu}{r(\mathbf{m})} \tag{16}$$

A scalar function $r(\mathbf{m})$ that satisfies this equation exists if and only if vector $\dot{\mathbf{m}}^* + \mathbf{P}_m\omega$ is parallel to vector $\mathbf{m} \times \nu$ for all $\mathbf{m} \in \Omega(S)$, that is,

$$(\dot{\mathbf{m}}^* + \mathbf{P}_m\omega) \times (\mathbf{m} \times \nu) = \mathbf{0} \tag{17}$$

If we note that $(\dot{\mathbf{m}}^*, \mathbf{m}) = 0$ and $(\mathbf{P}_m\omega, \mathbf{m}) = 0$, equation (17) reduces to

$$[(\dot{\mathbf{m}}^*, \nu) + (\mathbf{P}_m\omega, \nu)]\mathbf{m} = \mathbf{0} \tag{18}$$

From equation (10), we see that the condition for this to hold for all $\mathbf{m} \in \Omega(S)$ is given by equation (15). □

Mathematical equivalents of equation (15) have been derived by many researchers [2, 17, 32]. It can be shown that equation (15) is simply the *infinitesimal limit* of the fundamental constraint of finite motion known as the *epipolar equation*, on which the finite motion analysis is based [16, 28, 31] (appendix B). Hence, we call equation (15) the *epipolar equation*, too.

The geometrical meaning of the epipolar equation for finite motion is that each feature point in the scene and the two viewpoints before and after the motion be coplanar (appendix B). The same meaning is attached to equation (15). If $\omega = \mathbf{0}$, equation (15) becomes $(\dot{\mathbf{m}}^*, \nu) = 0$, or $|\mathbf{m}, \dot{\mathbf{m}}, \nu| = 0$, which states that $\mathbf{m}$, $\dot{\mathbf{m}}$, and ν be coplanar. The second term $(\mathbf{m}, \mathbf{K}\mathbf{m})$ compensates for the effect of camera rotation.

The epipolar equation for finite motion is written in terms of the "essential matrix" (appendix B), to which is assigned the *decomposability constraint* that its singular values be 1, 1, and 0 [8]. This constraint exactly corresponds to the decomposability constraint of equations (11).

5 Analytical Solution of 3-D Interpretation

Since the formulation is parallel to finite-motion analysis, the procedures developed for finite motion can be applied in exactly the same way. Note that the epipolar equation (15) is linear in the essential parameters $\{\mathbf{K}, \nu\}$, which have nine components (since $\mathbf{K}$ is symmetric). Hence, if the twisted flow is observed at at least "eight" pixels, we obtain the infinitesimal version the *eight-point algorithm* [16, 28]. This type of linearization was first proposed by Zhuang et al. [32]. Consider the optimization

$$J = \int_{\Omega(S)} [(\dot{\mathbf{m}}^*, \nu) + (\mathbf{m}, \mathbf{K}\mathbf{m})]^2 \, d\Omega(\mathbf{m}) \to \min \tag{19}$$

As in the case of finite motion, it is difficult to obtain an analytical solution that attains the exact minimum in the presence of noise, so we consider the linearized approach by regarding the essential parameters $\{\mathbf{K}, \nu\}$ as independent variables, ignoring the decomposability contraint (11) temporarily.

Since J is a quadratic form in nine essential parameters, it is minimized by the eigenvector of the nine-dimensional coefficient matrix for the smallest eigenvalue [32]. However, it seems more reasonable to find the minimum under the constraint that ν be a unit vector. Since J is a quadratic polynomial in $\mathbf{K}$, matrix $\mathbf{K}$ is analytically expressed as a linear form in ν. If it is substituted back, J becomes a quadratic form in ν, which is minimized by the unit eigenvector of the coefficient matrix for the smallest eigenvalue. Since the decomposability constraint (11) may not be satisfied, we estimate the rotation velocity ω by

$$\|\mathbf{K} - (\omega, \nu)\,\mathbf{I} + \frac{1}{2}(\omega\nu^{\top} + \nu\omega^{\top})\|^2 \to \min \tag{20}$$

The solution is given as follows (appendix C):

$$\omega = \frac{1}{2}[\mathrm{tr}\,\mathbf{K} + 3(\nu, \mathbf{K}\nu)]\,\nu - 2\mathbf{K}\nu \tag{21}$$

This approach was first proposed by Tagawa et al. [25, 26, 27].

In summary, our procedure is stated as follows. For a given twisted flow $\dot{\mathbf{m}}^*(\mathbf{m})$, define tensors $\mathcal{L} = (L_{ij})$, $\mathcal{M} = (M_{ijk})$, and $\mathcal{N} = (N_{ijkl})$ by

$$L_{ij} = \int_{\Omega(S)} \dot{m}_i^* \dot{m}_j^* \, d\Omega(\mathbf{m})$$

$$M_{ijk} = \int_{\Omega(S)} \dot{m}_i^* m_j m_k \, d\Omega(\mathbf{m})$$

$$N_{ijkl} = \int_{\Omega(S)} m_i m_j m_k m_l \, d\Omega(\mathbf{m}) \tag{22}$$

Let N_{ijkl}^{-1} be the inverse of N_{ijkl} defined by

$$\sum_{k,l=1}^{3} N_{ijkl}^{-1} N_{klmn} = \delta_{im}\,\delta_{jn} \tag{23}$$

where δ_{ij} is the Kronecker delta, taking value 1 if $i = j$ and 0 otherwise (see appendix D for the computational procedure).

Theorem 3: *The translation velocity $\nu = (\nu_i)$ is given as the unit eigenvector of the matrix $\mathbf{A} = (A_{ij})$ defined by*

$$A_{ij} = L_{ij} - \sum_{k,l,m,n=1}^{3} M_{ikl} N_{klmn}^{-1} M_{jmn} \tag{24}$$

for the smallest eigenvalue. The rotation velocity ω is given by

$$\omega = \frac{1}{2}[\mathrm{tr}\,\mathbf{K} + 3(\nu, \mathbf{K}\nu)]\,\nu - 2\mathbf{K}\nu \tag{25}$$

where the matrix $\mathbf{K} = (K_{ij})$ is defined by

$$K_{ij} = -\sum_{k,l,m=1}^{3} N_{ijkl}^{-1} M_{mkl}\,\nu_m \tag{26}$$

Proof: The function J defined by (19) is rewritten as

$$J = \sum_{i,j=1}^{3} \nu_i L_{ij} \nu_j + 2\sum_{i,j,k=1}^{3} \nu_i M_{ijk} K_{jk} + \sum_{i,j,k,l=1}^{3} K_{ij} N_{ijkl} K_{kl} \tag{27}$$

For a fixed $\nu = (\nu_i)$, this attains its minimum when $\partial J/\partial K_{ij} = 0$, which is written as

$$\sum_{k,l=1}^{3} N_{ijkl} K_{kl} = -\sum_{m=1}^{3} M_{mij}\,\nu_m \tag{28}$$

The solution is given by equation (26). Equations (26) and (28) reduce equation (27) to

$$J = \sum_{i,j=1}^{3} \nu_i \left(L_{ij} - \sum_{k,l,m,n=1}^{3} M_{ikl} N_{klmn}^{-1} M_{jmn} \right) \nu_j \tag{29}$$

If we define matrix $\mathbf{A} = (A_{ij})$ by equation (24), we can write $J = (\nu, \mathbf{A}\nu)$, which takes its minimum when ν is the unit eigenvector of $\mathbf{A}$ for the smallest eigenvalue. □

Note that tensors $\mathcal{L}$ and $\mathcal{M}$ are computed from the observed flow, whereas tensor $\mathcal{N}$ does not involve the flow: it is determined solely by the geometry of the field of view S. If the field of view S is small, the computation of the inverse $\mathcal{N}^{-1}$ may be unstable: the solution of the linear equation (28) may be sensitively affected by the noise in the tensor $\mathcal{M}$ on the right-hand side.

As is well known, the degree of numerical instability for solving simultaneous linear equations is measured by the *condition number* (= (the largest singular value)/(the smallest singular value)). For simplicity, assume a circular field of view S around the image origin o. Let α be the angle between the Z-axis and the outermost line of sight (figure 3a). Then, $\mathbf{m} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)^{\top}$ in spherical coordinates, and we can replace the integral $\int_{\Omega(S)} d\Omega(\mathbf{m})$ by $\int_0^{2\pi} d\phi \int_0^{\alpha} \sin\theta\, d\theta$. Figure 3b shows the condition number of $\mathcal{N}$ (i.e., of equation (28)) with respect to α. We can clearly see that the condition number rapidly

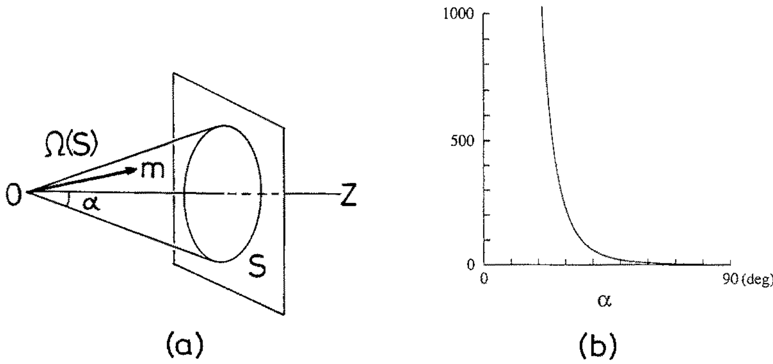

Fig. 3. (a) Circular field of view. (b) Condition number of $\mathcal{N}$.

increases as the angle α decreases. Thus, a wide field of view is desirable for stable 3-D interpretation.

6 Noise and Statistical Bias

Suppose the detected flow contains noise. The twisted flow observed at a pixel of N-vector $\mathbf{m}$ has the form $\dot{\mathbf{m}}^* = \bar{\dot{\mathbf{m}}}^* + \Delta\dot{\mathbf{m}}^*$, where $\bar{\dot{\mathbf{m}}}^*$ is the exact value in the absence of noise. By definition, $\dot{\mathbf{m}}^*$ and $\bar{\dot{\mathbf{m}}}^*$ are both orthogonal to $\mathbf{m}$. Hence, the error component $\Delta\dot{\mathbf{m}}^* = \dot{\mathbf{m}}^* - \bar{\dot{\mathbf{m}}}^*$ is also orthogonal to $\mathbf{m}$.

As a first approximation, let us assume that $\Delta\dot{\mathbf{m}}^*$ is an independent random variable whose distribution is orthogonally isotropic around $\mathbf{m}$ with a constant root mean square σ. This means that the error of optical flow is approximately $f\sigma$ in all orientations on the image plane in the sense of root mean square. From our model, we obtain

$$E[\Delta\dot{\mathbf{m}}^* \Delta\dot{\mathbf{m}}^{*\mathrm{T}}] = \frac{\sigma^2}{2}(\mathbf{I} - \mathbf{m}\mathbf{m}^{\mathrm{T}}) \tag{30}$$

where $E[\cdot]$ denotes expectation (see [11, 12] for detailed discussions on noise models).

From the definition of tensors $\mathcal{L} = (L_{ij})$ and $\mathcal{M} = (M_{ijk})$, their expectations are

$$E[L_{ij}] = \int_{\Omega(S)} E[\dot{m}_i^* \dot{m}_j^*]\, d\Omega(\mathbf{m})$$

$$E[M_{ijk}] = \int_{\Omega(S)} E[\dot{m}_i^*]\, m_j\, m_k\, d\Omega(\mathbf{m}) \tag{31}$$

while tensor $\mathcal{N} = (N_{ijkl})$ is not affected by noise. Since $E[\dot{\mathbf{m}}^*] = E[\bar{\dot{\mathbf{m}}}^* + \Delta\dot{\mathbf{m}}^*] = \bar{\dot{\mathbf{m}}}^*$ and

$$\begin{aligned} E[\dot{\mathbf{m}}^* \dot{\mathbf{m}}^{*\mathrm{T}}] &= E[(\bar{\dot{\mathbf{m}}}^* + \Delta\dot{\mathbf{m}}^*)(\bar{\dot{\mathbf{m}}}^* + \Delta\dot{\mathbf{m}}^*)^{\mathrm{T}}] \\ &= \bar{\dot{\mathbf{m}}}^* \bar{\dot{\mathbf{m}}}^{*\mathrm{T}} + E[\Delta\dot{\mathbf{m}}^* \Delta\dot{\mathbf{m}}^{*\mathrm{T}}] \\ &= \bar{\dot{\mathbf{m}}}^* \bar{\dot{\mathbf{m}}}^{*\mathrm{T}} + \frac{\sigma^2}{2}(\mathbf{I} - \mathbf{m}\mathbf{m}^{\mathrm{T}}) \end{aligned} \tag{32}$$

we obtain

$$E[\mathcal{L}] = \bar{\mathcal{L}} + \frac{\sigma^2}{2}(\Omega(S)\,\mathbf{I} - \mathbf{M}), \qquad E[\mathcal{M}] = \bar{\mathcal{M}} \tag{33}$$

where $\bar{\mathcal{L}}$ and $\bar{\mathcal{M}}$ are the exact values of tensors $\mathcal{L}$ and $\mathcal{M}$, respectively; $\mathbf{M}$ is the *moment matrix* defined by

$$\mathbf{M} = \int_{\Omega(S)} \mathbf{m}\mathbf{m}^{\mathrm{T}}\, d\Omega(\mathbf{m}) \tag{34}$$

Since $\mathcal{L}$ and $\mathcal{M}$ are defined by summation over a very large number of pixels, we can expect the $\mathcal{L} \approx E[\mathcal{L}]$ and $\mathcal{M} \approx E[\mathcal{M}]$ (the law of large number). Hence, from equation (24)

$$\mathbf{A} \approx \bar{\mathbf{A}} + \frac{\sigma^2}{2}[\Omega(S)\,\mathbf{I} - \mathbf{M}] \tag{35}$$

where $\bar{\mathbf{A}}$ is the exact value of $\mathbf{A}$ in the absence of noise. We can interpret this to be the expectation of $\mathbf{A}$. Thus, the matrix $\mathbf{A}$ is statistical bias

$$E[\Delta\mathbf{A}] = \frac{\sigma^2}{2}[\Omega(S)\,\mathbf{I} - \mathbf{M}] \tag{36}$$

Since the translation velocity $\boldsymbol{\nu}$ is the unit eigenvector of $\mathbf{A}$ for the smallest eigenvalue, it also has statistical bias $E[\Delta\boldsymbol{\nu}] = O(E[\Delta\mathbf{A}]) = O(\sigma^2)$ according to the well-known "perturbation theorem" (e.g., see [11]). Hence, the rotation velocity $\boldsymbol{\omega}$ computed from it has also statistical bias.

7 Renormalization

Define matrix $\hat{\mathbf{A}}$ by

$$\hat{\mathbf{A}} = \mathbf{A} - c[\Omega(S)\,\mathbf{I} - \mathbf{M}] \tag{37}$$

where $c = \sigma^2/2$. Then, the unit eigenvector of $\hat{\mathbf{A}}$ for the smallest eigenvalue is an unbiased estimator of the translation velocity $\boldsymbol{\nu}$. However, the constant c is unknown. Ideally, it should be chosen so that $E[\hat{\mathbf{A}}] = \bar{\mathbf{A}}$, but this is impossible unless the noise characteristics are known. On the other hand, if $E[\hat{\mathbf{A}}] = \bar{\mathbf{A}}$ and if $\bar{\boldsymbol{\nu}}$ is the exact translation velocity, we have the relationship

$$E[(\bar{\boldsymbol{\nu}}, \hat{\mathbf{A}}\bar{\boldsymbol{\nu}})] = (\bar{\boldsymbol{\nu}}, E[\hat{\mathbf{A}}]\bar{\boldsymbol{\nu}}) = (\bar{\boldsymbol{\nu}}, \bar{\mathbf{A}}\bar{\boldsymbol{\nu}}) = 0 \tag{38}$$

because J defined in (19) is rewritten as $J = (\boldsymbol{\nu}, \mathbf{A}\boldsymbol{\nu})$, which takes its absolute minimum 0 for the exact solution $\bar{\boldsymbol{\nu}}$. Hence, we seek a unit vector $\boldsymbol{\nu}$ and a constant c such that $(\boldsymbol{\nu}, \hat{\mathbf{A}}\boldsymbol{\nu}) = 0$ holds. If such $\boldsymbol{\nu}$ and c are not unique, we choose those for which c is minimum. This means

$$\frac{(\boldsymbol{\nu}, \mathbf{A}\boldsymbol{\nu})}{(\boldsymbol{\nu}, (\Omega(S)\,\mathbf{I} - \mathbf{M})\,\boldsymbol{\nu})} \rightarrow \min \tag{39}$$

Tagawa et al. [26, 27] pointed out that, if we put $\mathbf{B} = \Omega(S)\,\mathbf{I} - \mathbf{M}$, the solution is given by the unit (generalized) eigenvector problem $\mathbf{A}\boldsymbol{\nu} = c\mathbf{B}\boldsymbol{\nu}$, and the solution is easily obtained by computing the "square root" of the positive definite matrix $\mathbf{B}$ (appendix E).

Thus, we obtain the following procedure, which we call *renormalization*:

renormalization ($\dot{\mathbf{m}}^*$)

1. Compute tensors $\mathcal{L} = (L_{ij})$, $\mathcal{M} = (M_{ijk})$, and $\mathcal{N} = (N_{ijkl})$ defined by equations (22), and compute the inverse $\mathcal{N}^{-1} = (N^{-1}_{ijkl})$ of $\mathcal{N}$ (appendix D). Also, compute the moment matrix $\mathbf{M} = (M_{ij})$ given by equation (34).
2. Compute the matrixes $\mathbf{A} = (A_{ij})$ and $\mathbf{B} = (M_{ij})$ defined by

$$A_{ij} = L_{ij} - \sum_{k,l,m,n=1}^{3} M_{ikl}\, N^{-1}_{klmn}\, M_{jmn} \tag{40}$$

$$B_{ij} = \Omega(S)\, \delta_{ij} - M_{ij} \tag{41}$$

3. Let c be the unit (generalized) eigenvector for the smallest (generalized) eigenvalue of the generalized problem $\mathbf{A}\boldsymbol{\nu} = c\mathbf{B}\boldsymbol{\nu}$.
4. Compute matrix $\mathbf{K}$ by equation (26) and compute the rotation velocity $\boldsymbol{\omega}$ by equation (25). Return $\{\boldsymbol{\omega}, \boldsymbol{\nu}\}$.

From the above argument, we can also see that the (generalized) eigenvalue c measures the mean square noise level, which can also be veiwed as a *goodness of fit* of the computed solution. If $\mathbf{u}_1$ and $\mathbf{u}_2$ are the unit eigenvectors of $\hat{\mathbf{A}} = \mathbf{A} - c\mathbf{B}$ for the largest and intermediate eigenvalues λ_1 and λ_2, respectively, we can roughly interpret $\mathbf{u}_1$ and $\mathbf{u}_2$ as the orientations in which $\boldsymbol{\nu}$ is, respectively, the least and most likely to perturb, and λ_1 and λ_2 as indicating the degree of stability, although exact quantitative analysis is very difficult.

The renormalization technique shown here can be applied to many other problems including vanishing-point and focus of expansion estimation and conic fitting [13, 14]. In such problems, iterations are necessary for adjusting the weights for the least squares, but the convergence is proved to be quadratic [13, 14].

Example. Figure 4a shows a simulated optical flow defined over an image frame of 512×512 pixels with $f = 600$ (pixels). The camera is supposedly moving inside a rectangular room whose walls are tessellated with squares tiles. The corresponding twisted flow is shown in figure 4b. Random Gaussian noise with standard deviation 1 (pixel) is independently added to the x and y components of each flow velocity. The motion parameters $\{\boldsymbol{\omega}, \boldsymbol{\nu}\}$ are computed 100 times, each time using different noise. Figure 5 shows one example of 3-D reconstruction. The true shape is indicated by broken lines.

Figure 6 shows the error histograms of the motion parameters $\{\boldsymbol{\omega}, \boldsymbol{\nu}\}$ without renormalization—figure 6a is the histogram of the difference in angular velocity (in deg/s) between the true rotation velocity $\bar{\boldsymbol{\omega}}$ and the computed $\boldsymbol{\omega}$; figure 6b is the histogram of the difference in orientation (in deg) between the true translation velocity $\bar{\boldsymbol{\nu}}$ and the computed $\boldsymbol{\nu}$. We can clearly observe the existence of statistical bias. Figure 7 shows the corresponding results computed by applying renormalization. We can see that the bias has diminished.

8 Discussions

8.1 Reinterpretation of Renormalization

If $\dot{\mathbf{m}}$ (or equivalently $\bar{\dot{\mathbf{m}}}^*$) is the source of noise, as we assumed in our analysis, the most reasonable approach from a statistical point of view is to minimize $\int_{\Omega(S)} \|\Delta\dot{\mathbf{m}}\|^2\, d\Omega(\mathbf{m})$. Since $\Delta\dot{\mathbf{m}} = \dot{\mathbf{m}} - \bar{\dot{\mathbf{m}}}$ and the true flow $\bar{\dot{\mathbf{m}}}$ has the form of equation (4), we have

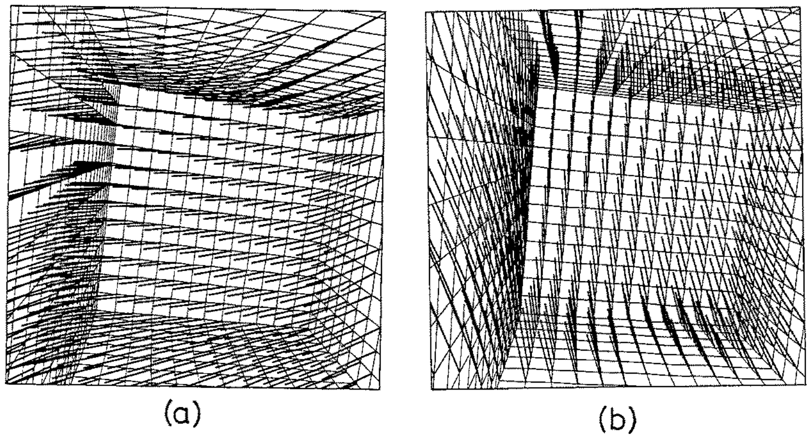

Fig. 4. (a) Optical flow. (b) Twisted flow.

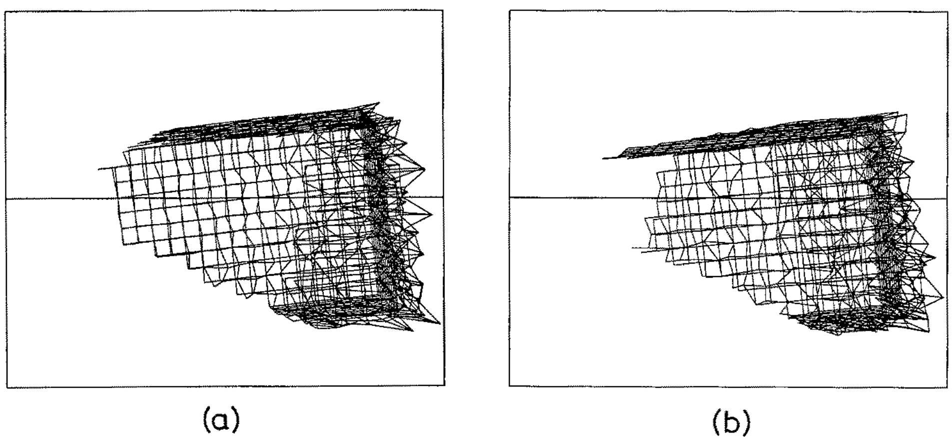

Fig. 5. An example of 3-D reconstruction from the flow of figure 4: (a) Side view; (b) Top view. The broken lines indicate the true shape.

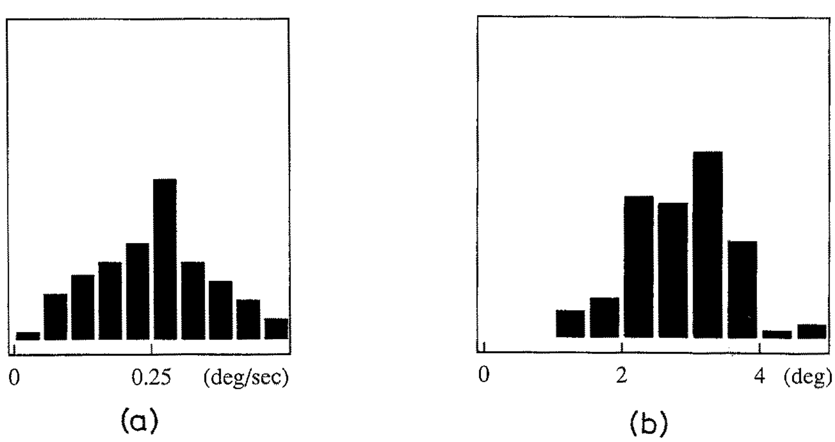

Fig. 6. (a) Histogram of error in rotation velocity without normalization. (b) Histogram of error in translation velocity without renormalization.

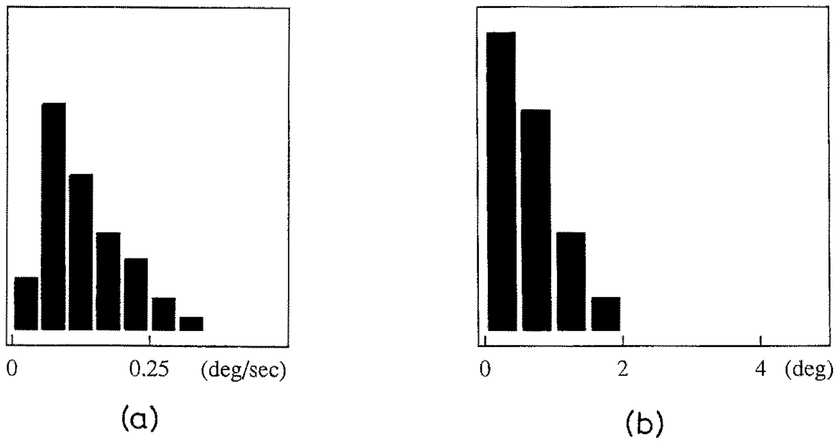

Fig. 7. (a) Histogram of error in rotation velocity with renormalization. (b) Histogram of error in translation velocity with renormalization.

$$J_0 = \int_{\Omega(S)} \left\| \dot{\mathbf{m}} + \omega \times \mathbf{m} + \frac{\mathbf{P}_{\mathrm{m}}\nu}{r(\mathbf{m})} \right\|^2 d\Omega(\mathbf{m}) \rightarrow \min \tag{42}$$

If the motion parallax equation (6) is substituted, it is easy to see (appendix F) that (42) reduces to

$$J_1 = \int_{\Omega(S)} \frac{[(\dot{\mathbf{m}}^*, \nu) + (\mathbf{m}, \mathbf{K}\mathbf{m})]^2}{(\nu, \mathbf{P}_{\mathrm{m}}\nu)} d\Omega(\mathbf{m}) \rightarrow \min \tag{43}$$

As pointed out by Tagawa et al. [26], our renormalization technique can be viewed as replacing J_1 as the following J_1':

$$J_1' = \frac{\int_{\Omega(S)} [(\dot{\mathbf{m}}^*, \nu) + (\mathbf{m}, \mathbf{K}\mathbf{m})]^2 \, d\Omega(\mathbf{m})}{\int_{\Omega(S)} (\nu, \mathbf{P}_{\mathrm{m}}\nu) \, d\Omega(\mathbf{m})} \rightarrow \min \tag{44}$$

If the essential parameters $\{\nu, \mathbf{K}\}$ are regarded as independent (i.e., linearized) variables and the matrixes **A** and **B** are defined as in section 7, the above minimization reduces to

$$J_2 = \frac{(\nu, \mathbf{A}\nu)}{(\nu, \mathbf{B}\nu)} \rightarrow \min \tag{45}$$

because $\int_{\Omega(S)} \mathbf{P}_{\mathrm{m}} \, d\Omega(\mathbf{m}) = \mathbf{B}$.

This explains why the solution is unbiased. Since $\dot{\mathbf{m}} \neq \bar{\dot{\mathbf{m}}}$, the function J_1' is not necessarily minimized by the true motion parameters $\{\bar{\omega}, \bar{\nu}\}$. Substituting $\dot{\mathbf{m}}^* = \bar{\dot{\mathbf{m}}}^* + \Delta\dot{\mathbf{m}}^*$, we obtain

$$J_1' = \frac{\int_{\Omega(S)}[(\nu, (\Delta\dot{\mathbf{m}}^*\Delta\dot{\mathbf{m}}^{*\mathrm{T}})\nu) + (\Delta\dot{\mathbf{m}}, \ldots) + \ldots] \, d\Omega(\mathbf{m})}{\int_{\Omega(S)}(\nu, \mathbf{P}_{\mathrm{m}}\nu) \, d\Omega(\mathbf{m})} \tag{46}$$

From equation (30) and $E[\Delta\dot{\mathbf{m}}) = 0$, we obtain

$$E[J_1'] = \frac{\int_{\Omega(S)}[\sigma^2(\nu, \mathbf{P}_{\mathrm{m}}\nu)/2 + ((\bar{\dot{\mathbf{m}}}^*, \nu) + (\mathbf{m}, \mathbf{K}\mathbf{m}))^2] \, d\Omega(\mathbf{m})}{\int_{\Omega(S)}(\nu, \mathbf{P}_{\mathrm{m}}\nu) \, d\Omega(\mathbf{m})}$$
$$= \frac{\sigma^2}{2} + \frac{\int_{\Omega(S)}[(\bar{\dot{\mathbf{m}}}^*, \nu) + (\mathbf{m}, \mathbf{K}\mathbf{m})]^2 \, d\Omega(\mathbf{m})}{(\nu, \mathbf{B}\nu)} \tag{47}$$

which is evidently minimized by the true motion parameters $\{\bar{\omega}, \bar{\nu}\}$. Thus, J_1' is minimized by the true solution *in expectation.* This solves the unsolved bias issue discussed by Daniilidis and Nagel [3].

8.2 The Critical Surface

Our analytical procedure yields a unique solution as long as the essential parameters $\{\mathbf{K}, \nu\}$ are uniquely determined. However, there exists a special type of object surface such that the epipolar equation (15) admits multiple solutions: it has the form (see appendix G)

$$(\mathbf{r}, \tilde{\mathbf{K}}\mathbf{r}) = (\nu \times \tilde{\nu}, \mathbf{r}) \tag{48}$$

where $\tilde{\mathbf{K}}$ is a symmetric matrix and $\tilde{\nu}$ is a unit vector such that $\{\tilde{\mathbf{K}}, \tilde{\nu}\} \neq \{\mathbf{K}, \nu\}$. Let us call this surface the *weak critical surface.* This is generally a quadric surface passing through the viewpoint *O.*

However, if all spurious essential parameters are not decomposable, the solution is theoretically unique (but difficult to compute numerically). On the other hand, if equations (15) are satisfied by some spurious essential parameters that are *decomposable,* the 3-D interpretation is inherently ambiguous, and the surface is called the (*strong*) *critical surface.*

By definition, the (strong) critical surface is also the weak critical surface, but the converse does not necessarily hold. Geometric properties of the critical surface have been studied in detail [6, 11, 17, 20]. For example, it is known (see appendix G) that:

- The (strong) critical surface is generally a ruled quadric surface, and hence is a hyperboloid of one sheet or its degeneracy.
- The (strong) critical surface for translation velocity $\boldsymbol{\nu}$ and spurious essential parameters $\{\tilde{\mathbf{K}}, \tilde{\boldsymbol{\nu}}\}$ degenerates to two planes, one passing through the viewpoint O, if and only if the rotation velocity $\tilde{\omega}$ defined by $\{\tilde{\mathbf{K}}, \tilde{\boldsymbol{\nu}}\}$ is orthogonal to $\boldsymbol{\nu}$ and $\tilde{\boldsymbol{\nu}}$. The two planes are orthogonal to each other.

8.3 Image Plane Representation

So far, all equations are written in terms of the N-vector $\mathbf{m}$ $(= N[(x, y, f)^T])$ and the N-velocity $\dot{\mathbf{m}}$. This means that the image is represented on an image sphere of unit radius centered at the viewpoint. On the other hand, many researchers favor the "image plane representation," in which an image point is represented by the position vector $\mathbf{x} = (x, y, f)^T$ and image velocity is represented by $\dot{\mathbf{x}}$. However, the image sphere representation and the image plane representation are mathematically equivalent, since $\mathbf{m}$ and $\mathbf{x}$ have the same information: they both indicate the direction of the line of sight.

Here, we translate some of our equations into the image plane representation for the sake of comparison. Let $\mathbf{k} = (0, 0, 1)^T$. Instead of the depth $r(\mathbf{m})$, let us use $Z(\mathbf{x})$ defined by $r(\mathbf{m})\mathbf{m} = Z(\mathbf{x})\mathbf{x}$ or $Z(\mathbf{x}) = r(\mathbf{m})/\|\mathbf{x}\|$. This means that the distance of the surface point corresponding to $\mathbf{x}$ from the XY plane is $fZ(\mathbf{x})$.

The position vector $\mathbf{x}$ is related to the N-vector $\mathbf{m}$ in the form

$$\mathbf{x} = \frac{f\mathbf{m}}{(\mathbf{k}, \mathbf{m})} \tag{49}$$

from which we obtain

$$\dot{\mathbf{x}} = \|\mathbf{x}\|\mathbf{Q}_x\dot{\mathbf{m}} \tag{50}$$

where we define

$$\mathbf{Q}_x = \mathbf{I} - \frac{1}{f}\mathbf{x}\mathbf{k}^T \tag{51}$$

Note that $\mathbf{Q}_x\mathbf{a} = -\mathbf{k} \times (\mathbf{x} \times \mathbf{a})/f$ for any vector $\mathbf{a}$. Applying equation (50) to equation (4) and noting that $\mathbf{x} = \|\mathbf{x}\|\mathbf{m}$ and $Z(\mathbf{x}) = r(\mathbf{m})/\|\mathbf{x}\|$, we obtain

$$\dot{\mathbf{x}} = -\mathbf{Q}_1(\omega \times \mathbf{x}) - \frac{\mathbf{Q}_1\mathbf{P}_1\boldsymbol{\nu}}{Z(\mathbf{x})} \tag{52}$$

Noting the identities $\mathbf{Q}_x(\omega \times \mathbf{x}) = -\|\mathbf{x}\|^2\,\mathbf{k} \times \mathbf{P}_m\omega/f$ and $\mathbf{Q}_x\mathbf{P}_m = \mathbf{Q}_x$, we obtain the flow equation in the form

$$\dot{\mathbf{x}} = \frac{\|\mathbf{x}\|^2}{f}\mathbf{k} \times \mathbf{P}_m\omega - \frac{\mathbf{Q}_x\boldsymbol{\nu}}{Z(\mathbf{x})} \tag{53}$$

The motion parallax equation (6) is replaced by

$$Z(\mathbf{x}) = \frac{\|\mathbf{Q}_x\boldsymbol{\nu}\|^2}{\|\mathbf{x}\|^2|\mathbf{k}, \mathbf{P}_m\omega, \mathbf{Q}_x\boldsymbol{\nu}|/f - (\dot{\mathbf{x}}, \mathbf{Q}_x\boldsymbol{\nu})} \tag{54}$$

If we define the "twisted flow" by $\dot{\mathbf{x}}^* = \mathbf{x} \times \dot{\mathbf{x}}$ and note that $\mathbf{x} \times (\mathbf{k} \times \mathbf{P}_m\omega) = -f\,\mathbf{P}_m\omega$ and $\mathbf{x} \times \mathbf{Q}_x\boldsymbol{\nu} = \mathbf{x} \times \boldsymbol{\nu}$, equation (13) becomes

$$\dot{\mathbf{x}}^* = -\|\mathbf{x}\|^2\,\mathbf{P}_m\omega - \frac{\mathbf{x} \times \boldsymbol{\nu}}{Z(\mathbf{x})} \tag{55}$$

The epipolar equation (15) reads

$$(\dot{\mathbf{x}}^*, \boldsymbol{\nu}) + (\mathbf{x}, \mathbf{K}\mathbf{x}) = 0 \tag{56}$$

Hence, the analytical procedure described in section 5 holds if $\dot{\mathbf{m}}^*$, $\mathbf{m}$, and $\int_{\Omega(S)} d\Omega(\mathbf{m})$ are replaced by $\dot{\mathbf{x}}^*$, $\mathbf{x}$, and $\int_S dx\,dy$, respectively.

The optimization (42) is replaced by

$$I_0 = \int_S \left\| \dot{\mathbf{x}} - \frac{\|\mathbf{x}\|^2}{f}\mathbf{k} \times \mathbf{P}_m\omega + \frac{\mathbf{Q}_x\boldsymbol{\nu}}{Z(\mathbf{x})} \right\|^2 dx\,dy \to \min \tag{57}$$

If this is minimized with respect to the depth $Z(\mathbf{x})$ by using equation (54), this becomes as follows (appendix H):

$$I_1 = \int_S \frac{[(\dot{\mathbf{x}}^*, \boldsymbol{\nu}) + (\mathbf{x}, \mathbf{K}\mathbf{x})]^2}{f^2(\boldsymbol{\nu}, \mathbf{Q}_x^T\mathbf{Q}_x\boldsymbol{\nu})} dx\,dy \to \min \tag{58}$$

If this is replaced by

$$I_1' = \frac{\int_S [(\dot{\mathbf{x}}^*, \boldsymbol{\nu}) + (\mathbf{x}, \mathbf{K}\mathbf{x})]^2\,dx\,dy}{\int_S f^2(\boldsymbol{\nu}, \mathbf{Q}_x^T\mathbf{Q}_x\boldsymbol{\nu})\,dx\,dy} dx\,dy \to \min \tag{59}$$

The solution is still unbiased by the same reason we showed before. If we "linearize" this by regarding the essential parameters $\{\boldsymbol{\nu}, \mathbf{K}\}$ as independent variables, the problem reduces to

$$I_2 = \frac{(\boldsymbol{\nu}, \tilde{\mathbf{A}}\boldsymbol{\nu})}{f^2(\boldsymbol{\nu}, \tilde{\mathbf{B}}\boldsymbol{\nu})} \to \min \tag{60}$$

where $\tilde{\mathbf{A}}$ is defined by equation (24) after replacing $\dot{\mathbf{m}}^*$, $\mathbf{m}$, and $\int_{\Omega(S)} d\Omega(\mathbf{m})$ by $\dot{\mathbf{x}}^*$, $\mathbf{x}$, and $\int_S dx\, dy$, respectively, and

$$\tilde{\mathbf{B}} = \int_S \mathbf{Q}_x^\top \mathbf{Q}_x \, dx\, dy$$

$$= \int_S \begin{pmatrix} 1 & & \frac{-x}{f} \\ & 1 & \frac{-y}{f} \\ \frac{-x}{f} & \frac{-y}{f} & \frac{x^2 + y^2}{f} \end{pmatrix} dx\, dy \quad (61)$$

The same result can be obtained by error analysis in exactly the same way as in section 6 if we use the noise model

$$E[\Delta\dot{\mathbf{x}}\, \Delta\dot{\mathbf{x}}^\top] = \frac{f^2\, \sigma^2(\mathbf{I} - \mathbf{k}\mathbf{k}^\top)}{2}$$

but we omit the details.

9 Concluding Remarks

This article has studied 3-D interpretation of optical flow induced by a general camera motion relative to a surface of general shape. First, we described, by using the "image sphere representation," an analytical procedure that yields an exact solution when the data are exact: we solved the *epipolar equation* written in terms of the *essential parameters* and the *twisted optical flow*. Introducing a simple model of noise, we have shown that the solution is shown to be statistically biased. Generalizing the approach of Tagawa et al. [25, 26, 27], we presented an algorithm called *renormalization*, which automatically adjusts to unknown image noise. A random-number simulation was given to observe its effectiveness. A brief discussion was given to the mathematical structure of renormalization, the critical surface, which is concisely described in terms of the essential parameters introduced in this article, and the use of the image plane representation.

Although our renormalization procedure solves the computational problem of 3-D interpretation, it does not have a very practical value unless optical flow is fairly accurately detected from real images. So, the importance of accurate optical-flow detection could not be emphasized too much.

Acknowledgments

The author thanks Toshio Iwasaki of Toshiba Corporation for doing numerical experiments when he was at Gunma University and Norio Tagawa of Tokyo Metropolitan University for various discussions on this research.

References

1. G. Adiv, Determining three-dimensional motion and structure from optical flow generated by several moving objects, *IEEE Trans. Patt. Anal. Mach. Intell.* 7:384–401, 1985.
2. A.R. Bruss and B.K.P. Horn, Passive navigation, *Comput. Vis. Graph. Image Process.* 21:3–20, 1983.
3. K. Daniilidis and H.-H. Nagel, Analytical results on error sensitivity of motion estimation from two views, *Image Vis. Comput.* 8:287–303, 1990.
4. L.S. Dreschler and H.-H. Nagel, Volumetric model and 3-D trajectory of a moving car dervied from monocular TV-frame sequences of a street scene, *Comput. Graph. Image Process.* 20:199–228, 1982.
5. D.J. Heeger and A.D. Jepson, Subspace methods for recovering rigid motion I: Algorithm and implementation, *Intern. J. Comput. Vis.* 7:95–117, 1992.
6. B.K.P. Horn, Motion fields are hardly ever ambiguous, *Intern. J. Comput. Vis.* 1:259–274, 1987.
7. B.K.P. Horn, Relative orientation, *Intern. J. Comput. Vis.* 4:59–78, 1990.
8. T.S. Huang and O.D. Faugeras, Some properties of the *E* matrix in two-view motion estimation, *IEEE Trans. Patt. Anal. Mach. Intell.* 11:1310–1312, 1989.
9. K. Kanatani, *Group-Theoretical Methods in Image Understanding*, Springer: Berlin, 1990.
10. K. Kanatani, Computational projective geometry, *Comput. Vis. Graph. Image Process.* 54:333–348, 1991.
11. K. Kanatani, *Geometric Computation for Machine Vision*, Oxford University Press: Oxford, 1993.
12. K. Kanatani, Unbiased estimation and statistical analysis of 3-D rigid motion from two views, *IEEE Trans. Patt. Anal. Mach. Intell.* 15:37–50, 1993.
13. K. Kanatani, Renormalization for unbiased estimation, *Proc. 4th Intern. Conf. Comput. Vis.*, May 1993, Berlin, pp. 599–606.
14. K. Kanatani, Statistical bias for conic fitting and renormalization, *IEEE Trans. Patt. Anal. Mach. Intell.* (to appear).
15. H.C. Longuet-Higgins and K. Prazdny, The interpretation of a moving retinal image, *Proc. Roy. Soc. London* B-208,: 385–397, 1980.
16. H.C. Longuet-Higgins, A computer algorithm for reconstructing a scene from two projections, *Nature* 293(10):133–135, 1981.
17. S. Maybank, The angular velocity associated with the optical flowfield arising from motion through rigid environment, *Proc. Roy. Soc. London. A* 401:317–326, 1985.
18. H.C. Longuet-Higgins, Multiple interpretations of a pair of images of a surface, *Proc. Roy. Soc. London A* 418:1–15, 1988.

19. S.J. Maybank, Ambiguity in reconstruction from image correspondences, *Image Vis. Comput.* 9:93–99, 1991.
20. S. Negahdaripour, Critical surface pairs and triplets, *Intern. J. Comput. Vis.* 3:293–312, 1989.
21. S. Negahdaripour, Multiple interpretation of the shape and motion of objects from two perspective images, *IEEE Trans. Patt. Anal. Mach. Intell.* 12:1025–1030, 1990.
22. K. Prazdny, Determining the instantaneous direction of motion from optical flow generated by a curvilinearly moving observer, *Comput. Graph. Image Process.* 17:238–248, 1981.
23. J.H. Reiger and D.T. Lawton, Processing differential image motion, *J. Opt. Soc. Amer. A* 2:165–175, 1984.
24. M. Subbarao, Interpretation of image flow: Rigid curved surfaces in motion, *Intern. J. Comput. Vis.* 2:77–96, 1988.
25. N. Tagawa, T. Toriu, and T. Endoh, Closed method for reconstructing the 3-dimensional rigid motion, *IEICE Technical Report*, IE90-108/PRU90-139, 1991, pp. 23–29 (in Japanese).
26. N. Tagawa, T. Toriu, T. Endoh, Y. Ikeda, and T. Moriya, Optimization of 3-D rigid motion estimation from optical flow, *IEICE Technical Report*, IE92-136/PRU92-159, 1993, pp. 25–32 (in Japanese).
27. N. Tagawa, T. Toriu, and T. Endoh, Un-biased linear algorithm for recovering three-dimensional motion from optical flow, *IEICE Trans. Infor. Sys.* (to appear).
28. R.Y. Tsai and T.S. Huang, Uniqueness and estimation for three-dimensional motion parameters of rigid objects with curved surfaces, *IEEE Trans. Patt. Anal. Mach. Intell.* 6:13–27, 1984.
29. Y. Yasumoto and G. Medioni, Robust estimation of three-dimensional motion parameters from a sequence of image frames using regularization, *IEEE Trans. Patt. Anal. Mach. Intell.* 8:464–471, 1986.
30. A.M. Waxman, B. Kamgar-Parsi, and M. Subbarao, Closed-form solutions to image flow equations for 3D structure and motion, *Intern. J. Comput. Vis.* 1:239–258, 1987.
31. J. Weng, T.S. Huang, and N. Ahuja, Motion and structure from two perspective views: Algorithms, error analysis, and error estimation, *IEEE Trans. Patt. Anal. Mach. Intell.* 11:451–467, 1989.
32. X. Zhuang, T.S. Huang, N. Ahuja, and R.M. Haralick, A simplified linear optical flow-motion algorithm, *Comput. Vis. Graphics Image Process.* 21:3–20, 1983.

Appendix A: Test for Pure Rotation

Theoretically, $\boldsymbol{\nu} = \mathbf{0}$ only when the flow field $\dot{\mathbf{m}}(\mathbf{m})$ has the form

$$\dot{\mathbf{m}} = \mathbf{m} \times \boldsymbol{\omega} \tag{62}$$

and the depth $r(\mathbf{m})$ is indeterminate: no 3-D information is available other than $\boldsymbol{\omega}$. In the presence of noise, we can fit a vector $\boldsymbol{\omega}$ such that

$$\int_{\Omega(S)} \|\dot{\mathbf{m}} - \mathbf{m} \times \boldsymbol{\omega}\|^2 \, d\Omega(\mathbf{m}) \to \min \tag{63}$$

Then, we test if the computed $\boldsymbol{\omega}$ satisfies

$$\|\dot{\mathbf{m}} - \mathbf{m} \times \boldsymbol{\omega}\| < \epsilon \qquad \mathbf{m} \in \Omega(S) \tag{64}$$

for an appropriately set error tolerance ϵ. The solution of (63) is given as follows (we omit the derivation; see [11]):

$$[\Omega(S)\,\mathbf{I} - \mathbf{M}]\,\boldsymbol{\omega} = \mathbf{b} \tag{65}$$

$$\mathbf{M} = \int_{\Omega(S)} \mathbf{m}\mathbf{m}^{\mathrm{T}} \, d\Omega(\mathbf{m})$$

$$\mathbf{b} = \int_{\Omega(S)} \dot{\mathbf{m}} \times \mathbf{m} \, d\Omega(\mathbf{m}) \tag{66}$$

Appendix B: Epipolar Equation for Finite Motion

If the camera is rotated around the center of the lens by $\mathbf{R}$ (rotation matrix) and translated by $\mathbf{h}$, a point in the scene having N-vector $\mathbf{m}$ and depth r moves to a point with N-vector $\mathbf{m}'$ and depth r' relative to the camera (figure B.1a). From figure B.1b, we have

$$\mathbf{h} = r\mathbf{m} - r'\,\mathbf{R}\mathbf{m}' \tag{67}$$

Note that vector $\mathbf{m}'$ in the second frame is rotated by $\mathbf{R}$. The depths r and r' satisfying equation (67) exist if and only if the three vectors $\mathbf{m}$, $\mathbf{h}$, and $\mathbf{R}\mathbf{m}'$ are coplanar, that is

$$|\mathbf{m}, \mathbf{h}, \mathbf{R}\mathbf{m}'| = 0 \tag{68}$$

This equation, often called the *epipolar equation*, was first derived by Longuet-Higgins [16] and Tsai and Huang [28]. Let $\mathbf{r}_1$, $\mathbf{r}_2$, and $\mathbf{r}_3$ be the three columns of $\mathbf{R}$, that is, $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$. If we introduce the notation $\mathbf{h} \times \mathbf{R} = (\mathbf{h} \times \mathbf{r}_1, \mathbf{h} \times \mathbf{r}_2, \mathbf{h} \times \mathbf{r}_3)$ and define matrix $\mathbf{G} = \mathbf{h} \times \mathbf{R}$, the above epipolar equation is written as

$$(\mathbf{m}, \mathbf{G}\mathbf{m}') = 0 \tag{69}$$

The matrix $\mathbf{G}$ is called the *essential matrix* [28].

If the camera motion is very small, we can write

$$\mathbf{h} = \boldsymbol{\nu}\Delta t + O(\Delta t^2)$$

$$\mathbf{R} = \mathbf{I} + \boldsymbol{\omega} \times \mathbf{I}\Delta t + O(\Delta t^2) \tag{70}$$

Hence,

$$\mathbf{G} = \boldsymbol{\nu} \times \mathbf{I}\Delta t + \boldsymbol{\nu} \times (\boldsymbol{\omega} \times \mathbf{I})\,\Delta t^2 + O(\Delta t_2) \tag{71}$$

If this and $\mathbf{m}' = \mathbf{m} + \dot{\mathbf{m}}\Delta t + O(\Delta t^2)$ are substituted, then the epipolar equation (69) becomes

$$([\mathbf{m}, \boldsymbol{\nu} \times (\boldsymbol{\omega} \times \mathbf{m})] + (\mathbf{m}, \boldsymbol{\nu} \times \dot{\mathbf{m}}))\,\Delta t^2 + O(\Delta t^3) = 0 \tag{72}$$

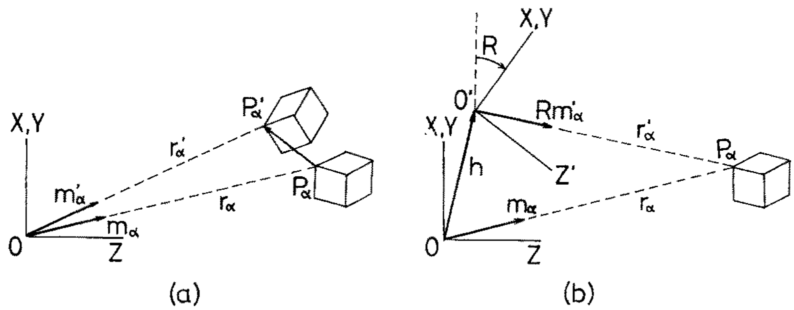

Fig. B.1. Camera motion $\{\mathbf{R}, \mathbf{h}\}$ relative to a point P in the scene. (a) Description with respect to the camera coordinate system. (b) Description with respect to the scene coordinate system.

Dividing this by Δt^2 and taking the limit as $\Delta t \to 0$ and noting that

$$(\mathbf{m}, \boldsymbol{\nu} \times (\boldsymbol{\omega} \times \mathbf{m})) = (\mathbf{m}, (\boldsymbol{\nu}, \mathbf{m})\boldsymbol{\omega} - (\boldsymbol{\nu}, \boldsymbol{\omega})\mathbf{m}) = -(\mathbf{m}, \mathbf{K}\mathbf{m}) \tag{73}$$

$$(\mathbf{m}, \boldsymbol{\nu} \times \dot{\mathbf{m}}) = -(\mathbf{m} \times \dot{\mathbf{m}}, \boldsymbol{\nu}) = -(\dot{\mathbf{m}}^*, \boldsymbol{\nu}) \tag{74}$$

we obtain

$$(\dot{\mathbf{m}}^*, \boldsymbol{\nu}) + (\mathbf{m}, \mathbf{K}\mathbf{m}) = 0 \tag{75}$$

which is the epipolar equation for optical flow.

Appendix C: Proof of Equation (21)

In elements, the left-hand side of (20) is

$$J = \sum_{i,j=1}^{3} \left\{ K_{ij} - \sum_{k=1}^{3} \left[\delta_{ij}\nu_k - \frac{1}{2}(\delta_{ik}\nu_j + \nu_i\delta_{jk}) \right] \omega_k \right\}^2 \tag{76}$$

The condition $\partial J/\partial \omega_i = 0$ is written as

$$\sum_{j=1}^{3} (3\nu_i\nu_j + \delta_{ij})\omega_j = 2 \left(\sum_{j=1}^{3} K_{jj}\nu_i - \sum_{j=1}^{3} K_{ij}\nu_j \right) \tag{77}$$

or

$$(\mathbf{I} + 3\boldsymbol{\nu}\boldsymbol{\nu}^{\mathrm{T}})\boldsymbol{\omega} = 2(\mathrm{tr}\,\mathbf{K})\boldsymbol{\nu} - 2\mathbf{K}\boldsymbol{\nu} \tag{78}$$

Using the identity $(\mathbf{I} + 3\boldsymbol{\nu}\boldsymbol{\nu}^{\mathrm{T}})^{-1} = \mathbf{I} - 3\boldsymbol{\nu}\boldsymbol{\nu}^{\mathrm{T}}/4$, we obtain

$$\boldsymbol{\omega} = 2\left(\mathbf{I} - \frac{3}{4}\boldsymbol{\nu}\boldsymbol{\nu}^{\mathrm{T}} \right) [(\mathrm{tr}\,\mathbf{K})\boldsymbol{\nu} - \mathbf{K}\boldsymbol{\nu}] = \frac{1}{2}[\mathrm{tr}\,\mathbf{K} + 3(\boldsymbol{\nu}\,\mathbf{K}\boldsymbol{\nu})]\,\boldsymbol{\nu} - 2\mathbf{K}\boldsymbol{\nu} \tag{79}$$

Appendix D: Inverse of Tensor $\mathcal{N} = (N_{ijkl})$

Tensor $\mathcal{N} = (N_{ijkl})$ is symmetric with respect to its indexes. It defines a linear mapping from a symmetric matrix $\mathbf{A} = (A_{ij})$ to a symmetric matrix $\mathbf{B} = (B_{ij})$ in the form

$$B_{ij} = \sum_{k,l=1}^{3} N_{ijkl} A_{kl}$$

or symbolically $\mathbf{B} = \mathcal{N}\mathbf{A}$. Since a symmetric matrix has six independent elements, it can be identified with a six-dimensional vector. Identify symmetric matrixes $\mathbf{A} = (A_{ij})$ and $\mathbf{B} = (B_{ij})$ with the six-dimensional vectors

$$\begin{pmatrix} A_{11} \\ \sqrt{2}A_{12} \\ A_{22} \\ \sqrt{2}A_{13} \\ \sqrt{2}A_{23} \\ A_{33} \end{pmatrix}, \quad \begin{pmatrix} B_{11} \\ \sqrt{2}B_{12} \\ B_{22} \\ \sqrt{2}B_{13} \\ \sqrt{2}B_{23} \\ B_{33} \end{pmatrix} \tag{80}$$

respectively. If we identify the tensor $\mathcal{N}$ with the six-dimensional matrix

$$\begin{pmatrix} N_{1111} & \sqrt{2}N_{1112} & N_{1122} & \sqrt{2}N_{1113} & \sqrt{2}N_{1123} & N_{1133} \\ \sqrt{2}N_{1211} & 2N_{1212} & \sqrt{2}N_{1222} & 2N_{1213} & 2N_{1223} & \sqrt{2}N_{1233} \\ N_{2211} & \sqrt{2}N_{2212} & N_{2222} & \sqrt{2}N_{2213} & \sqrt{2}N_{2223} & N_{2233} \\ \sqrt{2}N_{1311} & 2N_{1312} & \sqrt{2}N_{1322} & 2N_{1313} & 2N_{1323} & \sqrt{2}N_{1333} \\ \sqrt{2}N_{2311} & 2N_{2312} & \sqrt{2}N_{2322} & 2N_{2313} & 2N_{2323} & \sqrt{2}N_{2333} \\ N_{3311} & \sqrt{2}N_{3312} & N_{3322} & \sqrt{2}N_{3313} & \sqrt{2}N_{3323} & N_{3333} \end{pmatrix} \tag{81}$$

the relationship $\mathbf{B} = \mathcal{N}\mathbf{A}$ can be viewed as either a transformation of matrix $\mathbf{A}$ into $\mathbf{B}$ by tensor $\mathcal{N}$ or a transformation of vector $\mathbf{A}$ into vector $\mathbf{B}$ by matrix $\mathcal{N}$. Adopting the second viewpoint, we can compute the inverse $\mathcal{N}^{-1}$ of the above six-dimensional matrix. Then, we have $\mathbf{A} = \mathcal{N}^{-1}\mathbf{B}$, which reads $A_{ij} = \Sigma_{i,j,k,l=1}^{3} N_{ijkl}^{-1} B_{kl}$ if $\mathcal{N}^{-1}$ is regarded as a tensor $\mathcal{N}^{-1} = (N_{ijkl}^{-1})$ by the same rule. This means that the tensor N_{ijkl}^{-1} thus defined is the inverse of N_{ijkl}.

Appendix E: Generalized Eigenvalue Problem

The matrix $\mathbf{B}$ is positive definite unless $\Omega(S)$ is of measure 0. In fact, for any nonzero vector $\mathbf{u}$, we have

$$(\mathbf{u}, \mathbf{B}\mathbf{u}) = (\mathbf{u}, \int_{\Omega(S)} \mathbf{P}_{\mathrm{m}}\, d\Omega(\mathbf{m})\mathbf{u})$$

$$= \int_{\Omega(S)} (\mathbf{u}, \mathbf{P}_{\mathrm{m}}\mathbf{u})\, d\Omega(\mathbf{m})$$

$$= \int_{\Omega(S)} \|\mathbf{P}_{\mathrm{m}}\mathbf{u}\|^2\, d\Omega(\mathbf{m}) > 0 \quad (82)$$

Let $\lambda_1 \geq \lambda_2 \geq \lambda_3$ (> 0) be the eigenvalues of $\mathbf{B}$, and $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ be the orthonormal system of the corresponding eigenvectors. Then, matrix $\mathbf{B}$ is expressed in the following form (*spectral decomposition* [11]):

$$\mathbf{B} = \lambda_1\mathbf{u}_1\mathbf{u}_1^{\mathrm{T}} + \lambda_2\mathbf{u}_2\mathbf{u}_2^{\mathrm{T}} + \lambda_3\mathbf{u}_3\mathbf{u}_3^{\mathrm{T}} \quad (83)$$

Its square root is defined by

$$\sqrt{\mathbf{B}} = \sqrt{\lambda_1}\mathbf{u}_1\mathbf{u}_1^{\mathrm{T}} + \sqrt{\lambda_2}\mathbf{u}_2\mathbf{u}_2^{\mathrm{T}} + \sqrt{\lambda_3}\mathbf{u}_3\mathbf{u}_3^{\mathrm{T}} \quad (84)$$

It is easy to confirm that $(\sqrt{\mathbf{B}})^2 = \mathbf{I}$. The generalized eigenvalue problem

$$\mathbf{A}\boldsymbol{\nu} = c\mathbf{B}\boldsymbol{\nu} \quad (85)$$

is rewritten as

$$\mathbf{A}(\sqrt{\mathbf{B}})^{-1}\sqrt{\mathbf{B}}\boldsymbol{\nu} = c\sqrt{\mathbf{B}}\sqrt{\mathbf{B}}\boldsymbol{\nu} \quad (86)$$

Multiplying $(\sqrt{\mathbf{B}})^{-1}$ from left on both sides and putting $\mathbf{u} = \sqrt{\mathbf{B}}\boldsymbol{\nu}$, we obtain

$$(\sqrt{\mathbf{B}})^{-1}\mathbf{A}(\sqrt{\mathbf{B}})^{-1}\mathbf{u} = c\mathbf{u} \quad (87)$$

Thus, the generalized eigenvalue problem (85) reduces to the orindary eigenvalue problem $\tilde{\mathbf{A}}\mathbf{u} = c\mathbf{u}$ for the symmetric matrix $\tilde{\mathbf{A}} = (\sqrt{\mathbf{B}})^{-1}\mathbf{A}(\sqrt{\mathbf{B}})^{-1}$, and $\boldsymbol{\nu} = (\sqrt{\mathbf{B}})^{-1}\mathbf{u}$. Note that

$$(\sqrt{\mathbf{B}})^{-1} = \frac{1}{\sqrt{\lambda_1}}\mathbf{u}_1\mathbf{u}_1^{\mathrm{T}} + \frac{1}{\sqrt{\lambda_2}}\mathbf{u}_2\mathbf{u}_2^{\mathrm{T}} + \frac{1}{\sqrt{\lambda_3}}\mathbf{u}_3\mathbf{u}_3^{\mathrm{T}} \quad (88)$$

Appendix F: Derivation of Equation (43)

For two vectors $\mathbf{a}$ and $\mathbf{b}$ ($\neq \mathbf{0}$), we have the following identity:

$$\|\mathbf{a} + t\mathbf{b}\|^2 \geq \frac{\|\mathbf{a} \times \mathbf{b}\|^2}{\|\mathbf{b}\|^2} \quad (89)$$

Equality holds for

$$t = -\frac{(\mathbf{a}, \mathbf{b})}{\|\mathbf{b}\|^2} \quad (90)$$

Hence,

$$\left\|\dot{\mathbf{m}} + \omega \times \mathbf{m} + \frac{\mathbf{P}_{\mathrm{m}}\boldsymbol{\nu}}{r(\mathbf{m})}\right\|^2 \geq \frac{\|(\dot{\mathbf{m}} + \omega \times \mathbf{m}) \times \mathbf{P}_{\mathrm{m}}\boldsymbol{\nu}\|^2}{\|\mathbf{P}_{\mathrm{m}}\boldsymbol{\nu}\|^2} \quad (91)$$

and equality holds for the motion parallax equation (6). Since $\dot{\mathbf{m}} = \dot{\mathbf{m}}^* \times \mathbf{m}$, we see that

$$(\dot{\mathbf{m}} + \omega \times \mathbf{m}) \times \mathbf{P}_{\mathrm{m}}\boldsymbol{\nu} = (\dot{\mathbf{m}}^* \times \mathbf{m} + \omega \times \mathbf{m}) \times \mathbf{P}_{\mathrm{m}}\boldsymbol{\nu}$$

$$= ((\dot{\mathbf{m}}^* + \omega) \times \mathbf{m}) \times \mathbf{P}_{\mathrm{m}}\boldsymbol{\nu}$$

$$= (\dot{\mathbf{m}}^* + \omega, \mathbf{P}_{\mathrm{m}}\boldsymbol{\nu})\mathbf{m}$$

$$= ((\dot{\mathbf{m}}^*, \mathbf{P}_{\mathrm{m}}\boldsymbol{\nu}) + (\omega, \mathbf{P}_{\mathrm{m}}\boldsymbol{\nu}))\mathbf{m}. \quad (92)$$

Since $\dot{\mathbf{m}}^*$ is orthogonal to $\mathbf{m}$, we have $(\dot{\mathbf{m}}^*, \mathbf{P}_{\mathrm{m}}\boldsymbol{\nu}) = (\dot{\mathbf{m}}^*, \boldsymbol{\nu})$. We also have $(\omega, \mathbf{P}_{\mathrm{m}}\boldsymbol{\nu}) = (\mathbf{m}, \mathbf{K}\mathbf{m})$ from equation (10). Thus,

$$\|(\dot{\mathbf{m}} + \omega \times \mathbf{m}) \times \mathbf{P}_{\mathrm{m}}\boldsymbol{\nu}\|^2 = [(\dot{\mathbf{m}}^*, \boldsymbol{\nu}) + (\mathbf{m}, \mathbf{K}\mathbf{m})]^2 \quad (93)$$

Since $\|\mathbf{P}_{\mathrm{m}}\boldsymbol{\nu}\|^2 = (\boldsymbol{\nu}, \mathbf{P}_{\mathrm{m}}\boldsymbol{\nu})$, we obtain (43) by integrating (91) over $\Omega(S)$. □

Appendix G: Critical Surface of Optical Flow

In order to analyze ambiguities *the rotation velocity* ω *need not be considered*, since rotational motion around the viewpoint O does not add any 3-D information (see appendix A), just as finite camera rotation does not affect the 3-D interpretation of finite motion. In other words, it is geometrically evident that if a surface yields spurious motion parameters $\{\tilde{\omega}, \tilde{\boldsymbol{\nu}}\}$ together with the true motion parameters $\{\mathbf{0}, \boldsymbol{\nu}\}$, the same surface must yield spurious motion parameters $\{\tilde{\omega} + \omega, \tilde{\boldsymbol{\nu}}\}$ if the true motion parameters are $\{\omega, \boldsymbol{\nu}\}$. Hence we can assume $\omega = \mathbf{0}$ without losing generality, although a spurious interpretation $\{\tilde{\omega}, \tilde{\boldsymbol{\nu}}\}$ such that $\tilde{\omega} \neq \mathbf{0}$ may be obtained. From proposition 2 and our assumption that $\omega = \mathbf{0}$, the twisted flow is given by

$$\dot{\mathbf{m}}^* = -\frac{\mathbf{m} \times \boldsymbol{\nu}}{r(\mathbf{m})} \quad (94)$$

A spurious solution $\{\tilde{\mathbf{K}}, \tilde{\boldsymbol{\nu}}\}$ satisfies the epipolar equation (15) if and only if

$$\left(-\frac{\mathbf{m} \times \nu}{r(\mathbf{m})}, \tilde{\nu} \right) + (\mathbf{m}, \tilde{\mathbf{K}}\mathbf{m}) = 0 \quad (95)$$

Putting $\mathbf{r} = r(\mathbf{m})\mathbf{m}$, we obtain the following results:

Theorem G.1: *The weak critical surface has the form*

$$(\mathbf{r}, \tilde{\mathbf{K}}\mathbf{r}) = (\nu \times \tilde{\nu}, \mathbf{r}) \quad (96)$$

where $\tilde{\mathbf{K}}$ is a symmetric matrix and $\tilde{\nu}$ is a unit vector such that $\{\tilde{\mathbf{K}}, \tilde{\nu}\} \neq \{\mathbf{K}, \nu\}$.

From this and the definition of the (strong) critical surface, we obtain

Theorem G.2: *The (strong) critical surface has the form*

$$(\mathbf{r}, \tilde{\mathbf{K}}\mathbf{r}) = (\nu \times \tilde{\nu}, \mathbf{r}) \quad (97)$$

where the pair $\{\tilde{\mathbf{K}}, \tilde{\nu}\}$ is decomposable and $\{\tilde{\mathbf{K}}, \tilde{\nu}\} \neq \{\mathbf{K}, \nu\}$.

Corollary G.1: *The (weak or strong) critical surface is generally a quadric surface passing through the viewpoint O.*

As the term "generally" implies, equation (97) does not necessarily describe a surface for special values of $\{\tilde{\mathbf{K}}, \tilde{\nu}\}$. For example:

— If $\tilde{\mathbf{K}} = \mathbf{O}$, the surface equation becomes $(\nu \times \tilde{\nu}, \mathbf{r}) = 0$, which describes a line passing through the viewpoint O.
— If $\tilde{\nu} = \pm\nu$, the surface equation becoms $(\mathbf{r}, \tilde{\mathbf{K}}\mathbf{r}) = 0$, which describes a set of lines passing through the viewpoint O.

Hence, we assume in the following that the (strong) critical surface is defined by a decomposable pair $\{\tilde{\mathbf{K}}, \tilde{\nu}\}$ such that $\tilde{\mathbf{K}} \neq \mathbf{O}$ and $\tilde{\nu} \neq \mathbf{0}$. The following is a restatement of the fact pointed out by Horn [6]:

Theorem G.3: *The (strong) critical surface is generally a ruled quadric, and hence is a hyperboloid of one sheet or its degeneracy.*

It can be shown that the spurious interpretation of the critical surface is also a quadric surface of the same type.

Proposition G.1 *The spurious interpretation of the (strong) critical surface $(\mathbf{r}, \tilde{\mathbf{K}}\mathbf{r}) = (\nu \times \tilde{\nu}, \mathbf{r})$ is*

$$(\mathbf{r}, \tilde{\mathbf{K}}'\mathbf{r}) = (\nu \times \tilde{\nu}, \mathbf{r}) \quad (98)$$

$$\tilde{\mathbf{K}}' = (\tilde{\omega}, \nu)\,\mathbf{I} - \frac{1}{2}(\tilde{\omega}\nu^{\top} + \nu\tilde{\omega}^{\top}) \quad (99)$$

It is well known that an optical flow resulting from a motion relative to a planar surface in the scene yields two 3-D interpretations. This means that any single planar surface can be the (strong) critical surface. However, we have just proved that the critical surface is a quadric surface passing through the viewpoint O. This apparent inconsistency is resolved if and only if the critical surface degenerates to two planes S_0 and S_1 in such a way that one of them, say S_0, passes through O so that the viewer can view only S_1. The condition that this degeneracy occurs is exactly parallel to the case of finite motion [18, 20] and is given as follows:

Theorem G.4: *The (strong) critical surface for translation velocity ν and spurious essential parameters $\{\tilde{\mathbf{K}}, \tilde{\nu}\}$ degenerates to two planes, one passing through the viewpoint O, if and only if the rotation velocity $\tilde{\omega}$ defined by $\{\tilde{\mathbf{K}}, \tilde{\nu}\}$ is orthogonal to ν and $\tilde{\nu}$.*

Since the two planes $(\tilde{\omega}, \mathbf{r}) = 0$ and $(\tilde{\nu}, \mathbf{r}) + c = 0$ have their surface normals $\tilde{\omega}$ and $\tilde{\nu}$, respectively, we find that:

Corollary G.2: *The two planes into which the critical surface degenerates are orthogonal to each other.*

Finally, we can easily confirm that the critical surface equation (97) is obtained by taking the infinitesimal limit of the critical surface equation for finite motion. Indeed, if the notations of appendix B are used, it can be shown (e.g., see [11]) that the critical surface equation for finite motion has the form

$$(\mathbf{r}, \tilde{\mathbf{G}}\mathbf{r}) = (\mathbf{r}, \tilde{\mathbf{G}}\mathbf{h}) \quad (100)$$

where $\tilde{\mathbf{G}}$ is a spurious essential matrix. For a small translation $\mathbf{h} = \nu\Delta t + O(\Delta t^2)$, the above critical surface equation is written as

$$(\mathbf{r}, \tilde{\mathbf{G}}\mathbf{r}) = (\mathbf{r}, \tilde{\mathbf{G}}\nu)\,[\Delta t + O(\Delta t^2)] \quad (101)$$

If the spurious motion parameters are infinitesimal, equation (71) also holds for spurious motion parameters, and we obtain

$$(\mathbf{r}, (\tilde{\nu}, \mathbf{r})\tilde{\omega} - (\tilde{\nu}, \tilde{\omega})\mathbf{r})\,\Delta t^2 + O(\Delta t^3)$$
$$= (\mathbf{r}, \tilde{\nu} \times \nu)\,\Delta t^2 + (\mathbf{r}, \tilde{\nu} \times (\tilde{\omega} \times \nu))\,\Delta t^3 + O(\Delta t^4) \quad (102)$$

Dividing both sides by Δt^2 and taking the limit as $\Delta t \to 0$, we obtain,

$$(\mathbf{r}, [\tilde{\omega}\tilde{\nu}^{\mathrm{T}} - (\tilde{\nu}, \tilde{\omega})\,\mathbf{I}]\,\mathbf{r}) = -(\nu \times \tilde{\nu}, \mathbf{r}) \tag{103}$$

The left-hand side is a symmetric quadratic form in $\mathbf{r}$ and can be written as $-(\mathbf{r}, \tilde{\mathbf{K}}\mathbf{r})$ if we define

$$\tilde{\mathbf{K}} = (\tilde{\nu}, \tilde{\omega})\,\mathbf{I} - \frac{1}{2}(\tilde{\omega}\tilde{\nu}^{\mathrm{T}} + \tilde{\nu}\tilde{\omega}^{\mathrm{T}}) \tag{104}$$

Hence, equation (103) is written as

$$(\mathbf{r}, \tilde{\mathbf{K}}\mathbf{r}) = (\nu \times \tilde{\nu}, \mathbf{r}) \tag{105}$$

which is the critical surface equation for optical flow.

Appendix H: Derivation of Equation (58)

From equation (89), we have

$$\|\dot{\mathbf{x}} - \frac{\|\mathbf{x}\|^2}{f}\mathbf{k} \times \mathbf{P}_{\mathrm{m}}\omega + \frac{\mathbf{Q}_{\mathrm{x}}\nu}{Z(\mathbf{x})}\|^2 \geq \frac{\|\dot{\mathbf{x}} \times Q_{\mathrm{x}}\nu - \|\mathbf{x}\|^2(\mathbf{k} \times \mathbf{P}_{\mathrm{m}}\omega) \times \mathbf{Q}_{\mathrm{x}}\nu/f\|^2}{\|\mathbf{Q}_{\mathrm{x}}\nu\|^2} \tag{106}$$

where equality holds for equation (54). It is easily shown that

$$\dot{\mathbf{x}} \times \mathbf{Q}_{\mathrm{x}}\nu = \frac{1}{f}(\mathbf{x} \times \dot{\mathbf{x}}, \nu)\mathbf{k} = \frac{1}{f}(\dot{\mathbf{x}}^*, \nu)\mathbf{k} \tag{107}$$

and

$$\begin{aligned}(\mathbf{k} \times \mathbf{P}_{\mathrm{m}}\omega) \times \mathbf{Q}_{\mathrm{x}}\nu &= -(\mathbf{P}_{\mathrm{m}}\omega, \mathbf{Q}_{\mathrm{x}}\nu)\mathbf{k} \\ &= -(\omega, \mathbf{P}_{\mathrm{m}}\mathbf{Q}_{\mathrm{x}}\nu)\mathbf{k} \\ &= -(\omega, \mathbf{P}_{\mathrm{m}}\nu)k \\ &= -(\mathbf{m}, \mathbf{K}_{\mathrm{m}})\mathbf{k} \\ &= -\frac{(\mathbf{x}, \mathbf{K}\mathbf{x})}{\|\mathbf{x}\|^2}\mathbf{k}\end{aligned} \tag{108}$$

Using these and integrating (106), we obtain (58). □