# Incorporating optical flow uncertainty information into a self-calibration procedure for a moving camera


M. J. Brooks[a], W. Chojnacki[a], A. Dick[a], A. van den Hengel[a], K. Kanatani[b], N. Ohta[b]

[a]Department of Computer Science, University of Adelaide, Adelaide, SA 5005, Australia

[b]Department of Computer Science, Gunma University, Kiryu, Gunma 376-8515, Japan



## ABSTRACT

In this paper we consider robust techniques for estimating structure from motion in the uncalibrated case. We show how information describing the uncertainty of the data may be incorporated into the formulation of the problem, and we explore the situations in which this appears to be advantageous. The structure recovery technique is based on a method for self-calibrating a single moving camera from instantaneous optical flow developed in previous work of some of the authors.[1] The method of self-calibration rests upon an equation that we term the differential epipolar equation for uncalibrated optical flow. This equation incorporates two matrices (analogous to the fundamental matrix in stereo vision) which encode information about the ego-motion and internal geometry of the camera. Any sufficiently large, non-degenerate optical flow field enables the ratio of the entries of the two matrices to be estimated. Under certain assumptions, the moving camera can be self-calibrated by means of closed-form expressions in the entries of these matrices. Reconstruction of the scene, up to a scalar factor, may then proceed using a straightforward method.[1] The critical step in this whole approach is therefore the accurate estimation of the aforementioned ratio. To this end, the problem is couched in a least-squares minimisation framework whereby candidate cost functions are derived via ordinary least squares, total least squares, and weighted least squares techniques. Various computational schemes are adopted for minimising the cost functions. Carefully devised synthetic experiments reveal that when the optical flow field is contaminated with inhomogeneous and anisotropic Gaussian noise, the best performer is the weighted least squares approach with renormalisation.



**Keywords:** Robustness, covariance matrix, self-calibration, egomotion, epipolar equation, uncalibrated camera


## 1. INTRODUCTION

Consider a perspective camera moving smoothly and freely through space amidst a stationary environment. At any time static elements of the 3D scene induce, via projection onto the camera's evolving image plane, an optical flow field. Such an instantaneous field is composed of *flow vectors* taking the form of pairs $(\boldsymbol{m}, \dot{\boldsymbol{m}})$, where $\boldsymbol{m}$ is the image location and $\dot{\boldsymbol{m}}$ is the image velocity of a specific element of the scene. It proves convenient to write $\boldsymbol{m}$ and $\dot{\boldsymbol{m}}$ as 3-vectors

$$\boldsymbol{m} = [m_1, m_2, 1]^T, \qquad \dot{\boldsymbol{m}} = [\dot{m}_1, \dot{m}_2, 0]^T,$$

where $(m_1, m_2)$ and $(\dot{m}_1, \dot{m}_2)$ represent, respectively, the element's image location and image velocity relative to a fixed image coordinate system. With this convention, the flow vectors $(\boldsymbol{m}, \dot{\boldsymbol{m}})$ satisfy the *differential epipolar equation for uncalibrated optical flow*[1,2]

$$\boldsymbol{m}^T \boldsymbol{W} \dot{\boldsymbol{m}} + \boldsymbol{m}^T \boldsymbol{C} \boldsymbol{m} = 0, \tag{1}$$

where $\boldsymbol{C}$ and $\boldsymbol{W}$ are $3 \times 3$ matrices encoding the information about the ego-motion and internal geometry of the camera that can directly be linked with the optical flow. The matrix $\boldsymbol{C}$ is symmetric, uniquely determined by the entries $c_{11}, c_{12}, c_{13}, c_{22}, c_{23}, c_{33}$, and the matrix $\boldsymbol{W}$ is skew-symmetric, determined by the entries $w_{12}, w_{13}, w_{23}$. It emerges that $\boldsymbol{C}$ and $\boldsymbol{W}$ are not independent, but are subject to the cubic constraint

$$\boldsymbol{w}^T \boldsymbol{C} \boldsymbol{w} = 0, \tag{2}$$

where $\boldsymbol{w} = [-w_{23}, w_{13}, -w_{12}]^T$. Both (1) and (2) are unaffected by multiplying $\boldsymbol{C}$ and $\boldsymbol{W}$ by a common scalar. As a consequence, the entire information concerning $\boldsymbol{C}$ and $\boldsymbol{W}$ extricable from optical flow resides in the composite ratio $(c_{11} : c_{12} : c_{13} : c_{22} : c_{23} : c_{33} : w_1 : w_2 : w_3)$, henceforth denoted $\boldsymbol{C} : \boldsymbol{W}$. Based on (1), the ratio $\boldsymbol{C} : \boldsymbol{W}$ may


*Authors' e-mail addresses:* {mjb,wojtek,anthony,hengel}@cs.adelaide.edu.au, {kanatani,ohta}@cs.gunma-u.ac.jp

uniquely be determined given at least eight optical flow vectors (given exactly seven flow vectors, either one or three ratios satisfy equations (1) and (2)).

As it turns out, structure recovery is essentially reduced to the determination of the ratio $\boldsymbol{C} : \boldsymbol{W}$.[1] In fact, the camera can be self-calibrated (under certain assumptions) by means of closed-form expressions evolved from this ratio.[1] Scene reconstruction, up to a scalar factor, can then be carried out by employing a procedure (also described in Ref. 1) based on the results of self-calibration and the optical flow data.

This paper is concerned with the practical problem of how the ratio $\boldsymbol{C} : \boldsymbol{W}$ might optimally be estimated from measurements of an optical flow field corrupted by random errors. A characteristic of our approach, in contrast with other work,[2–6] is that we deal with optical flow that is potentially sparse and whose elements may be subject to noise in both base location and velocity.

## 2. MODELLING NOISE IN OPTICAL FLOW DATA

Suppose that optical flow is computed at various points in the image, not necessarily densely across a grid. Let $\boldsymbol{m}_i = [m_{i1}, m_{i2}, 1]^T$ and $\dot{\boldsymbol{m}}_i = [\dot{m}_{i1}, \dot{m}_{i2}, 0]^T$ be the observed base location and observed velocity of the $i$th element of the optical flow field. Consider the system of equations obtained from (1) by substituting $\boldsymbol{m}_i$ and $\dot{\boldsymbol{m}}_i$ for $\boldsymbol{m}$ and $\dot{\boldsymbol{m}}$, respectively, for all $i$. If the number of the flow vectors $(\boldsymbol{m}_i, \dot{\boldsymbol{m}}_i)$ is precisely eight, then this system can be solved to obtain a unique ratio $\boldsymbol{C} : \boldsymbol{W}$. If the number of flow vectors exceeds eight, then this system is overdetermined and may have no solution. In this case, we are concerned with minimally adjusting the data to achieve a best fitting ratio $\boldsymbol{C} : \boldsymbol{W}$. A key step in the adjustment process is the adoption of an appropriate statistical model of the flow data. Here we assume a model falling into the category of *errors-in-variables* models.

Suppose that the observed location $\boldsymbol{m}_i$ is perturbed from a true location $\overline{\boldsymbol{m}}_i$ by noise represented by a random vector $\Delta\boldsymbol{m}_i = [\Delta m_{i1}, \Delta m_{i2}, 0]^T$. Likewise suppose that the observed velocity $\dot{\boldsymbol{m}}_i$ is perturbed from a true velocity $\overline{\dot{\boldsymbol{m}}}_i$ by noise represented by a random vector $\Delta\dot{\boldsymbol{m}}_i = [\Delta\dot{m}_{i1}, \Delta\dot{m}_{i2}, 0]^T$; we thus have

$$\boldsymbol{m}_i = \overline{\boldsymbol{m}}_i + \Delta\boldsymbol{m}_i \quad \text{and} \quad \dot{\boldsymbol{m}}_i = \overline{\dot{\boldsymbol{m}}}_i + \Delta\dot{\boldsymbol{m}}_i.$$

The true flow vectors $(\overline{\boldsymbol{m}}_i, \overline{\dot{\boldsymbol{m}}}_i)$ are assumed to satisfy equations (1) and (2) for some (unknown) matrices $\boldsymbol{C}$ and $\boldsymbol{W}$. In contrast, the observed flow vectors $(\boldsymbol{m}_i, \dot{\boldsymbol{m}}_i)$ will not in general satisfy equations (1) and (2) for any choice of $\boldsymbol{C}$ and $\boldsymbol{W}$. We set out to estimate the value of $\boldsymbol{C} : \boldsymbol{W}$ corresponding to the $(\overline{\boldsymbol{m}}_i, \overline{\dot{\boldsymbol{m}}}_i)$, based on the available data $(\boldsymbol{m}_i, \dot{\boldsymbol{m}}_i)$.

This task requires that we first characterise the uncertainty of our optical flow data. Assume that, for each $i$, $(\Delta m_{i1}, \Delta m_{i2})$ are drawn from a multivariate Gaussian distribution with zero mean value (the zero mean condition in the multivariate case meaning, of course, that $\mathsf{E}[\Delta m_{i1}] = \mathsf{E}[\Delta m_{i2}] = 0$, where $\mathsf{E}$ denotes the expected value), and assume that $(\Delta\dot{m}_{i1}, \Delta\dot{m}_{i2})$ are also drawn from a multivariate Gaussian distribution with zero mean value. Furthermore, assume that $\Delta m_{ik}$ and $\Delta\dot{m}_{il}$ are independent for any $i$ and any $k, l = 1, 2$, and that $\Delta m_{ik}$, $\Delta\dot{m}_{il}$, $\Delta m_{jp}$, $\Delta\dot{m}_{jq}$ are independent for $i \neq j$ and any $k, l, p, q = 1, 2$. The joint distribution of the $\Delta m_{i1}$ and $\Delta m_{i2}$ and that of the $\Delta\dot{m}_{i1}$ and $\Delta\dot{m}_{i2}$ are fully characterised by the *covariance matrices* of $\Delta\boldsymbol{m}_i$ and $\Delta\dot{\boldsymbol{m}}_i$, which alternatively can be viewed as the covariance matrices of $\boldsymbol{m}_i$ and $\dot{\boldsymbol{m}}_i$.

Recall that if $\boldsymbol{x} = [x_1, x_2, x_3]^T$ is a multivariate random variable perturbed from its mean value $\mathsf{E}[\boldsymbol{x}]$ by a random error $\Delta\boldsymbol{x} = [\Delta x_1, \Delta x_2, \Delta x_3]^T$ (so that $\boldsymbol{x} = \mathsf{E}[\boldsymbol{x}] + \Delta\boldsymbol{x}$), then the $3 \times 3$ covariance matrix $\mathrm{Cov}[\boldsymbol{x}]$ of $\boldsymbol{x}$ is defined by

$$(\mathrm{Cov}[\boldsymbol{x}])_{ij} = \mathsf{E}[\Delta x_i \Delta x_j].$$

The entry $(\mathrm{Cov}[\boldsymbol{x}])_{ij}$ describes the correlation between $x_i$ and $x_j$, and has the following properties:

- If $i = j$, then $(\mathrm{Cov}[\boldsymbol{x}])_{ij}$ is simply the variance of $x_i$.
- If $i \neq j$, then $(\mathrm{Cov}[\boldsymbol{x}])_{ij}$ is the covariance of $x_i$ and $x_j$; in particular, if $x_i$ and $x_j$ are independent, then $(\mathrm{Cov}[\boldsymbol{x}])_{ij} = 0$.
- $\mathrm{Cov}[\boldsymbol{x}]$ is symmetric.

In the case of the random vectors $\boldsymbol{m}_i$ and $\dot{\boldsymbol{m}}_i$, the third entries are constant. The absence of noise in these entries implies that all entries in the third row and the third column of the covariance matrices $\mathrm{Cov}\,[\boldsymbol{m}_i]$ and $\mathrm{Cov}\,[\dot{\boldsymbol{m}}_i]$ are zero. Thus each of these matrices takes the form

$$\mathrm{Cov}\,[\boldsymbol{x}] = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & 0 \\ \sigma_{12} & \sigma_2^2 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{3}$$

where, for $k = 1, 2$, $\sigma_k = (\mathsf{E}[\Delta x_k^2])^{1/2}$ is the standard deviation of $x_k$, and $\sigma_{12} = \mathsf{E}[\Delta x_1 \Delta x_2] = \mathsf{E}[\Delta x_2 \Delta x_1]$ is the covariance of $x_1$ and $x_2$.

Using covariance matrices we can model *homogeneous* or *inhomogeneous* noise, and *isotropic* or *anisotropic* noise. In a homogeneous noise model, the characteristics of the noise at each optical flow position are the same, as are those of the noise in each optical flow velocity observation. Thus $\mathrm{Cov}\,[\boldsymbol{m}_i]$ will be identical for all $i$, as will $\mathrm{Cov}\,[\dot{\boldsymbol{m}}_i]$. If $\mathrm{Cov}\,[\boldsymbol{m}_i] \neq \mathrm{Cov}\,[\boldsymbol{m}_j]$, or $\mathrm{Cov}\,[\dot{\boldsymbol{m}}_i] \neq \mathrm{Cov}\,[\dot{\boldsymbol{m}}_j]$ for some $i \neq j$, then the noise is inhomogeneous. In an isotropic noise model the noise variance does not depend on direction; that is, $\sigma_{12} = 0$ and $\sigma_1 = \sigma_2$ in each covariance matrix, the common variance possibly changing from one flow vector to another. If $\sigma_{12} \neq 0$ or $\sigma_1 \neq \sigma_2$ for some covariance matrix, then the noise is anisotropic. Figure 1 illustrates the properties of (in)homogeneous and (an)isotropic noise via various contours. We may regard each contour as a level set of the underlying normal distribution surface, with the characteristic that there is probability (say) 0.5 that an observed (noisy) value will occur within the region it encloses.

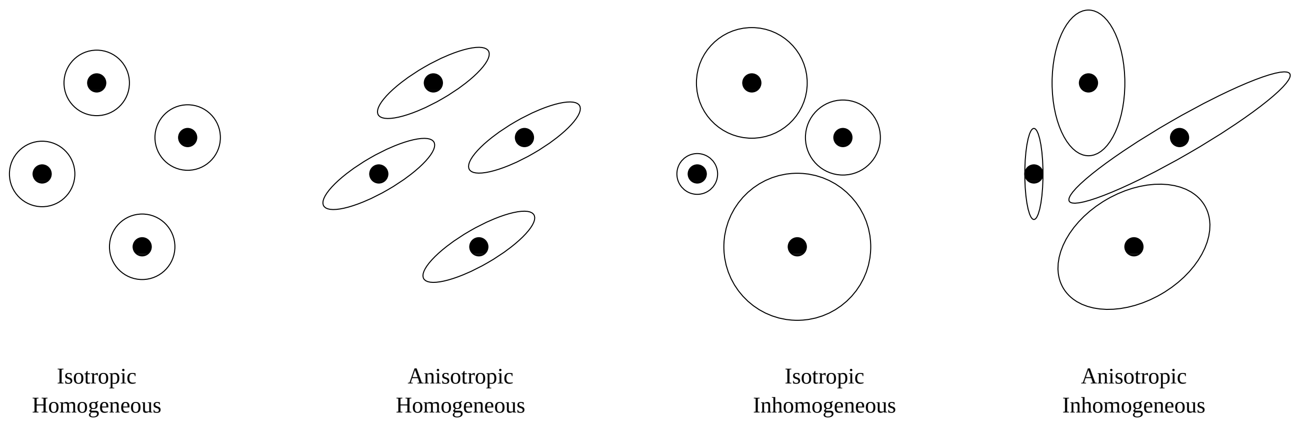


**Figure 1.** Example likelihood "contours" for (in)homogeneous and (an)isotropic noise models.

## 3. ESTIMATORS OF THE COMPOSITE RATIO

We now address the design of estimators of the composite ratio $\boldsymbol{C} : \boldsymbol{W}$. We develop three estimators minimising three different *cost functions* measuring the extent to which the measurements and candidate estimates fail to satisfy (1). The simplest of these cost functions will incorporate no prior information about the measurement errors. The other two, more involved functions will incorporate prior knowledge of measurement error statistics.

### 3.1. Ordinary least squares estimator

Given a single flow vector $(\boldsymbol{m}, \dot{\boldsymbol{m}})$ and a pair $(\boldsymbol{C}, \boldsymbol{W})$, let the corresponding *algebraic distance* or *algebraic residual* be defined as

$$R = \boldsymbol{m}^T \boldsymbol{W} \dot{\boldsymbol{m}} + \boldsymbol{m}^T \boldsymbol{C} \boldsymbol{m}.$$

Henceforth we shall only consider pairs $(\boldsymbol{C}, \boldsymbol{W})$ satisfying ${c_{11}}^2 + {c_{12}}^2 + {c_{13}}^2 + {c_{22}}^2 + {c_{23}}^2 + {c_{33}}^2 + {w_{12}}^2 + {w_{13}}^2 + {w_{23}}^2 = 1$. Under this assumption, the absolute value of $R$ depends on $\boldsymbol{C}$ and $\boldsymbol{W}$ exclusively through $\boldsymbol{C} : \boldsymbol{W}$ and can be viewed as a measure of the agreement between $(\boldsymbol{m}, \dot{\boldsymbol{m}})$ and $\boldsymbol{C} : \boldsymbol{W}$.

For each $i$, let $R_i$ be the algebraic distance corresponding to $(\boldsymbol{m}_i, \dot{\boldsymbol{m}}_i)$ and $\boldsymbol{C} : \boldsymbol{W}$; that is,

$$R_i = \boldsymbol{m}_i^T \boldsymbol{W} \dot{\boldsymbol{m}}_i + \boldsymbol{m}_i^T \boldsymbol{C} \boldsymbol{m}_i.$$

A straightforward estimate of $\boldsymbol{C} : \boldsymbol{W}$ is one minimising the cost function $\sum_i R_i^2$. Symbolically,

$$\langle \boldsymbol{C} : \boldsymbol{W} \rangle_{\mathrm{OLS}} = \arg \min_{\boldsymbol{C}:\boldsymbol{W}} \sum_i R_i^2, \tag{4}$$

where $\langle \boldsymbol{C} : \boldsymbol{W} \rangle_{\mathrm{OLS}}$ is the estimate of interest. We call it the *ordinary least squares* (OLS) *estimate*. The function assigning $\langle \boldsymbol{C} : \boldsymbol{W} \rangle_{\mathrm{OLS}}$ to a field of $(\boldsymbol{m}_i, \dot{\boldsymbol{m}}_i)$ will be termed the ordinary least squares *estimator*. The OLS estimator has the advantage of being amenable to a simple SVD-minimisation technique (specified later). We set aside the need to satisfy the cubic constraint. (Our initial experiments[7] indicate that little improvement results from a posterior correction of the ratio. One such correction is presented in Ref. 7. However, further study is needed to fully incorporate the cubic constraint into the estimation process.) We shall in due course use the OLS estimator as a base method against which to compare more sophisticated estimators, and in particular a technique that takes into account the nature of the uncertainty of the data.

### 3.2. Weighted least squares estimator

Of primary importance in estimator design is that different residuals $R_i$ may carry different statistical weight. When the $(\boldsymbol{m}_i, \dot{\boldsymbol{m}}_i)$ are treated as sample values of independent multivariate random variables, the $R_i$ are sample values of typically a *heteroscedastic* set of random variables. Recall that a set of random variables is said to be heteroscedastic if the member variables have different variances.

The larger the variance of a particular $R_i$, the less reliable this term is likely to be, and the more it should be devalued. Therefore, to account for heteroscedasticity, it is natural to replace the simple cost function $\sum_i R_i^2$ by the more complicated cost function $\sum_i R_i^2 / \mathrm{Var}[R_i]$. The corresponding *weighted least square* estimate $\langle \boldsymbol{C} : \boldsymbol{W} \rangle_{\mathrm{WLS}}$ is then given by

$$\langle \boldsymbol{C} : \boldsymbol{W} \rangle_{\mathrm{WLS}} = \arg \min_{\boldsymbol{C}:\boldsymbol{W}} \sum_i \frac{R_i^2}{\mathrm{Var}[R_i]}. \tag{5}$$

This paradigm has already been successfully used in a number of estimation problems in computer vision, notably in estimating fundamental matrices (see Ref. 8–11).

The variances $\mathrm{Var}[R_i]$ can readily be calculated. Given a scalar function $f$ of a multivariate argument $\boldsymbol{x}$, denote by $\partial_{\boldsymbol{x}} f$ the row vector of the corresponding partial derivatives

$$\partial_{\boldsymbol{x}} f = \left[ \frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right].$$

Given a 3-vector $\boldsymbol{x} = [x_1, x_2, x_3]^T$, designate by $\|\boldsymbol{x}\| = \sqrt{x_1^2 + x_2^2 + x_3^2}$ the Euclidean norm of $\boldsymbol{x}$. Expanding $R_i$ into a Taylor series in $(\boldsymbol{m}_i, \dot{\boldsymbol{m}}_i)$ about $(\overline{\boldsymbol{m}}_i, \overline{\dot{\boldsymbol{m}}}_i)$

$$R_i(\boldsymbol{m}_i, \dot{\boldsymbol{m}}_i) = R_i(\overline{\boldsymbol{m}}_i, \overline{\dot{\boldsymbol{m}}}_i) + \partial_{\boldsymbol{m}_i} R_i(\overline{\boldsymbol{m}}_i, \overline{\dot{\boldsymbol{m}}}_i) \Delta \boldsymbol{m}_i + \partial_{\dot{\boldsymbol{m}}_i} R_i(\overline{\boldsymbol{m}}_i, \overline{\dot{\boldsymbol{m}}}_i) \Delta \dot{\boldsymbol{m}}_i + O\left( (\|\Delta \boldsymbol{m}_i\| + \|\Delta \dot{\boldsymbol{m}}_i\|)^2 \right),$$

taking into account that $R_i(\overline{\boldsymbol{m}}_i, \overline{\dot{\boldsymbol{m}}}_i) = 0$ and

$$\partial_{\boldsymbol{m}_i} R_i(\overline{\boldsymbol{m}}_i, \overline{\dot{\boldsymbol{m}}}_i) = \partial_{\boldsymbol{m}_i} R_i(\boldsymbol{m}_i, \dot{\boldsymbol{m}}_i) + O\left(\|\Delta \boldsymbol{m}_i\| + \|\Delta \dot{\boldsymbol{m}}_i\|\right)$$

$$\partial_{\dot{\boldsymbol{m}}_i} R_i(\overline{\boldsymbol{m}}_i, \overline{\dot{\boldsymbol{m}}}_i) = \partial_{\dot{\boldsymbol{m}}_i} R_i(\boldsymbol{m}_i, \dot{\boldsymbol{m}}_i) + O\left(\|\Delta \boldsymbol{m}_i\| + \|\Delta \dot{\boldsymbol{m}}_i\|\right)$$

and ignoring terms of second or higher order in $\|\Delta \boldsymbol{m}_i\| + \|\Delta \dot{\boldsymbol{m}}_i\|$, we obtain the approximate equality

$$R_i = \partial_{\boldsymbol{m}_i} R_i \, \Delta \boldsymbol{m}_i + \partial_{\dot{\boldsymbol{m}}_i} R_i \, \Delta \dot{\boldsymbol{m}}_i$$

where the derivatives are evaluated at $(\boldsymbol{m}_i, \dot{\boldsymbol{m}}_i)$. The random vectors $\Delta \boldsymbol{m}_i$ and $\Delta \dot{\boldsymbol{m}}_i$ are independent, and so $\partial_{\boldsymbol{m}_i} R_i \, \Delta \boldsymbol{m}_i$ and $\partial_{\dot{\boldsymbol{m}}_i} R_i \, \Delta \dot{\boldsymbol{m}}_i$ are also independent, implying that

$$\mathrm{Var}[R_i] = \mathrm{Var}[\partial_{\boldsymbol{m}_i} R_i \, \Delta \boldsymbol{m}_i] + \mathrm{Var}[\partial_{\dot{\boldsymbol{m}}_i} R_i \, \Delta \dot{\boldsymbol{m}}_i].$$

Now

$$\mathrm{Var}\left[\partial_{\boldsymbol{m}_i} R_i\, \Delta \boldsymbol{m}_i\right] = \mathsf{E}\left[\partial_{\boldsymbol{m}_i} R_i\, \Delta \boldsymbol{m}_i \Delta \boldsymbol{m}_i{}^T\, (\partial_{\boldsymbol{m}_i} R_i)^T\right] = \partial_{\boldsymbol{m}_i} R_i\, \mathrm{Cov}\left[\boldsymbol{m}_i\right]\, (\partial_{\boldsymbol{m}_i} R_i)^T$$

and likewise

$$\mathrm{Var}\left[\partial_{\dot{\boldsymbol{m}}_i} R_i\, \Delta \dot{\boldsymbol{m}}_i\right] = \partial_{\dot{\boldsymbol{m}}_i} R_i\, \mathrm{Cov}\left[\dot{\boldsymbol{m}}_i\right]\, (\partial_{\dot{\boldsymbol{m}}_i} R_i)^T.$$

Therefore

$$\langle \boldsymbol{C} : \boldsymbol{W} \rangle_{\mathrm{WLS}} = \arg\min_{\boldsymbol{C}:\boldsymbol{W}} \sum_i \frac{R_i^2}{\partial_{\boldsymbol{m}_i} R_i\, \mathrm{Cov}\left[\boldsymbol{m}_i\right]\, (\partial_{\boldsymbol{m}_i} R_i)^T + \partial_{\dot{\boldsymbol{m}}_i} R_i\, \mathrm{Cov}\left[\dot{\boldsymbol{m}}_i\right]\, (\partial_{\dot{\boldsymbol{m}}_i} R_i)^T}. \tag{6}$$

Since

$$\partial_{\boldsymbol{m}_i} R_i = (2\boldsymbol{C}\boldsymbol{m}_i + \boldsymbol{W}\dot{\boldsymbol{m}}_i)^T \quad \text{and} \quad \partial_{\dot{\boldsymbol{m}}_i} R_i = -(\boldsymbol{W}\boldsymbol{m}_i)^T,$$

we finally obtain

$$\langle \boldsymbol{C} : \boldsymbol{W} \rangle_{\mathrm{WLS}} = \arg\min_{\boldsymbol{C}:\boldsymbol{W}} \sum_i \frac{(\boldsymbol{m}_i^T \boldsymbol{W} \dot{\boldsymbol{m}}_i + \boldsymbol{m}_i^T \boldsymbol{C} \boldsymbol{m}_i)^2}{(2\boldsymbol{C}\boldsymbol{m}_i + \boldsymbol{W}\dot{\boldsymbol{m}}_i)^T \mathrm{Cov}\left[\boldsymbol{m}_i\right] (2\boldsymbol{C}\boldsymbol{m}_i + \boldsymbol{W}\dot{\boldsymbol{m}}_i) + (\boldsymbol{W}\boldsymbol{m}_i)^T \mathrm{Cov}\left[\dot{\boldsymbol{m}}_i\right] \boldsymbol{W}\boldsymbol{m}_i}. \tag{7}$$

Note that the value of $\langle \boldsymbol{C} : \boldsymbol{W} \rangle_{\mathrm{WLS}}$ remains unchanged if the covariance matrices are multiplied by a common scalar.

Observe also that the weights (the denominators of fractions) entering the cost function in the right-hand side of (7) depend on parameters to be estimated. In fact many WLS estimators appearing in computer vision formulations operate with parameter-dependent weights.[9–11] In contrast, classical WLS estimators involve pure data weighting with no reference to parameters.

The above derivation of the WLS estimator follows the lines adopted in evolving an WLS estimator for the fundamental matrix.[9–11] Kanatani's theory of geometric fitting[3,12,13] provides an alternative and more foundational approach. Based on a statistical analysis of noise, it involves minimisation of the Mahalanobis distance between the observed and ideal flow data. It happens that, in this instance, the two approaches yield the same cost function.

### 3.3. Total least squares estimator

If the assumed noise is homogeneous and isotropic, and $\sigma$ is a common non-zero variance of significant entries of the noise vectors, then

$$\mathrm{Cov}\left[\boldsymbol{m}_i\right] = \mathrm{Cov}\left[\dot{\boldsymbol{m}}_i\right] = \begin{bmatrix} \sigma & 0 & 0 \\ 0 & \sigma & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

In this case the value defined by (6) will be called the *total least squares* (TLS) estimate of $\boldsymbol{C} : \boldsymbol{W}$. If we let $\| [x_1, x_2, x_3]^T \| = \sqrt{x_1^2 + x_2^2}$, then (6) becomes

$$\langle \boldsymbol{C} : \boldsymbol{W} \rangle_{\mathrm{TLS}} = \arg\min_{\boldsymbol{C}:\boldsymbol{W}} \sum_i \frac{R_i^2}{\|(\partial_{\boldsymbol{m}_i} R_i)^T\|^2 + \|(\partial_{\dot{\boldsymbol{m}}_i} R_i)^T\|^2}, \tag{8}$$

or in expanded form

$$\langle \boldsymbol{C} : \boldsymbol{W} \rangle_{\mathrm{TLS}} = \arg\min_{\boldsymbol{C}:\boldsymbol{W}} \sum_i \frac{(\boldsymbol{m}_i^T \boldsymbol{W} \dot{\boldsymbol{m}}_i + \boldsymbol{m}_i^T \boldsymbol{C} \boldsymbol{m}_i)^2}{\|2\boldsymbol{C}\boldsymbol{m}_i + \boldsymbol{W}\dot{\boldsymbol{m}}_i\|^2 + \|\boldsymbol{W}\boldsymbol{m}_i\|^2}.$$

Note that $\langle \boldsymbol{C} : \boldsymbol{W} \rangle_{\mathrm{TLS}}$ is independent of any particular (non-zero) value of $\sigma$.

## 4. COVARIANCE MATRIX GENERATION

Before we can experimentally compare the usefulness of the cost functions described in the previous section, we need to generate noisy synthetic optical flow data. This in turn requires that we synthesise or estimate appropriate covariance matrices. These issues we now consider.

### 4.1. Generating synthetic covariance matrices and data

The generation of noisy synthetic optical flow data is performed in three stages:

1. Generate ideal optical flow data $(\overline{\boldsymbol{m}}_i, \overline{\dot{\boldsymbol{m}}}_i)$ by means of a software simulation of a cloud of points viewed by a projective camera undergoing translation and rotation.

2. Determine a covariance matrix for each optical flow location and velocity as follows:

   (a) Set a range $[a, b]$ within which standard deviations will be confined in the next step.[†]

   (b) For each flow index $i$ and each $j = 1, 2$, select standard deviations $\sigma_{ij}$ and $\dot{\sigma}_{ij}$. If homogeneous noise is to be modelled, let
   $$\sigma_{ik} = \sigma_{jk} \quad \text{and} \quad \dot{\sigma}_{ik} = \dot{\sigma}_{jk} \quad \text{for all } i \neq j \text{ and all } k = 1, 2,$$
   and if isotropic noise is to be modelled, let
   $$\sigma_{i1} = \sigma_{i2} \quad \text{and} \quad \dot{\sigma}_{i1} = \dot{\sigma}_{i2} \quad \text{for all } i.$$
   Define two covariance matrices $\boldsymbol{\Sigma}_{\boldsymbol{m}_i}$ and $\boldsymbol{\Sigma}_{\dot{\boldsymbol{m}}_i}$ with principal components aligned to the $x$ and $y$ axes by setting
   $$\boldsymbol{\Sigma}_{\boldsymbol{m}_i} = \begin{bmatrix} \sigma_{i1}^2 & 0 & 0 \\ 0 & \sigma_{i2}^2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \boldsymbol{\Sigma}_{\dot{\boldsymbol{m}}_i} = \begin{bmatrix} \dot{\sigma}_{i1}^2 & 0 & 0 \\ 0 & \dot{\sigma}_{i2}^2 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

   (c) If anisotropic noise is to be modelled, for an angle $\phi \in [0, 2\pi)$ set
   $$\boldsymbol{O}_{\phi} = \begin{bmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$
   For each $i$, randomly select two angles $\phi_i$ and $\psi_i$, and define two covariance matrices with principal components rotated by $\phi_i$ and $\psi_i$, respectively, as follows:
   $$\operatorname{Cov}[\boldsymbol{m}_i] = \boldsymbol{O}_{\phi_i} \boldsymbol{\Sigma}_{\boldsymbol{m}_i} (\boldsymbol{O}_{\phi_i})^T \quad \text{and} \quad \operatorname{Cov}[\dot{\boldsymbol{m}}_i] = \boldsymbol{O}_{\psi_i} \boldsymbol{\Sigma}_{\dot{\boldsymbol{m}}_i} (\boldsymbol{O}_{\psi_i})^T.$$

3. Generate independent noises $\Delta\boldsymbol{m}_i$ and $\Delta\dot{\boldsymbol{m}}_i$ with associated covariance matrices $\operatorname{Cov}[\boldsymbol{m}_i]$ and $\operatorname{Cov}[\dot{\boldsymbol{m}}_i]$ determined in the previous step. Add $\Delta\boldsymbol{m}_i$ and $\Delta\dot{\boldsymbol{m}}_i$ to $\overline{\boldsymbol{m}}_i$ and $\overline{\dot{\boldsymbol{m}}}_i$ generated in Step 1 to obtain noisy flow data $\boldsymbol{m}_i$ and $\dot{\boldsymbol{m}}_i$, respectively.

## 5. PERFORMANCE ANALYSIS

### 5.1. Minimising the cost functions

The performance of our self-calibration technique depends in practice not only on the adopted cost function, but also on the means by which the cost is to be minimised.

Single eigenvalue decomposition offers a natural means of obtaining an OLS estimate of $\boldsymbol{C} : \boldsymbol{W}$. The key observation is that $\langle \boldsymbol{C} : \boldsymbol{W} \rangle_{\text{OLS}}$ can be identified with an eigenvector associated with the smallest eigenvalue of a certain matrix as follows. Noting that

$$\boldsymbol{m}^T \boldsymbol{W} \dot{\boldsymbol{m}} + \boldsymbol{m}^T \boldsymbol{C} \boldsymbol{m} = \boldsymbol{\theta}^T \boldsymbol{u}_{\boldsymbol{m},\dot{\boldsymbol{m}}},$$

where

$$\boldsymbol{\theta} = [c_{11}, c_{12}, c_{13}, c_{22}, c_{23}, c_{33}, w_{12}, w_{13}, w_{23}]^T \tag{9}$$

[†]This serves to ensure a necessary normalisation, as estimates are unaffected by common scaling of underlying covariance matrices.

and

$$\boldsymbol{u}_{\boldsymbol{m},\dot{\boldsymbol{m}}} = \begin{bmatrix} m_1^2 \\ 2m_1m_2 \\ 2m_1m_3 \\ m_2^2 \\ 2m_2m_3 \\ m_3^3 \\ m_1\dot{m}_2 - m_2\dot{m}_1 \\ m_1\dot{m}_3 - m_3\dot{m}_1 \\ m_2\dot{m}_3 - m_3\dot{m}_2 \end{bmatrix},$$

and subsequently rewriting $\sum_i R_i^2$ as $\boldsymbol{\theta}^T \boldsymbol{S}\boldsymbol{\theta}$, where $\boldsymbol{S} = \sum_i \boldsymbol{u}_{\boldsymbol{m}_i,\dot{\boldsymbol{m}}_i}(\boldsymbol{u}_{\boldsymbol{m}_i,\dot{\boldsymbol{m}}_i})^T$, it becomes obvious that $\langle \boldsymbol{C} : \boldsymbol{W}\rangle_{\mathrm{OLS}}$ is identifiable with the normalised vector $\widehat{\boldsymbol{\theta}}$ that minimises $\boldsymbol{\theta}^T \boldsymbol{S}\boldsymbol{\theta}$. Using Lagrange multipliers, it can be shown that $\widehat{\boldsymbol{\theta}}$ is an eigenvector of $\boldsymbol{S}$ corresponding to the least eigenvalue. This eigenvector can be robustly computed by performing a singular value decomposition on $\boldsymbol{S}$.

Because of the complex way in which $\boldsymbol{C}$ and $\boldsymbol{W}$ enter into the TLS and WLS cost functions, an iterative scheme is needed to minimise each of these functions. We experiment with three such schemes: Powell's method,[14] Sampson's method,[10,15] and Kanatani's renormalisation.[13] (Note that while being an outgrowth of the analysis of an iterative scheme designed for calculating TLS and WLS minimisers, renormalisation is not—strictly speaking—meant to calculate TLS and WLS minimisers but rather to generate estimates that on average are as good as TLS and WLS estimates.) Each of these methods is seeded with the OLS estimate. Minimising the TLS and WLS cost functions in three different ways results in possibly six different estimates of the ratio of interest.

## 5.2. Error Measures

In order to compare each combination of cost function and minimisation technique, we require a measure of the accuracy of a given estimate of $\boldsymbol{C} : \boldsymbol{W}$. Let $\boldsymbol{\theta}$ be an estimate of $\boldsymbol{C} : \boldsymbol{W}$ expressed as a normalised vector defined by (9), let $\boldsymbol{v} = [v_1, v_2, v_3]^T$ be an estimate of the camera's translational velocity (normalised), and let $\boldsymbol{\omega} = [\omega_1, \omega_2, \omega_3]^T$ be an estimate of the camera's rotational velocity. Let $\overline{\boldsymbol{\theta}}$, $\overline{\boldsymbol{v}}$, and $\overline{\boldsymbol{\omega}}$ be the respective true values. Assuming that we know $\overline{\boldsymbol{\theta}}$, $\overline{\boldsymbol{v}}$ and $\overline{\boldsymbol{\omega}}$, we can define three error measures:

- $\mathcal{E}_{\boldsymbol{\theta}} = \arccos(\boldsymbol{\theta} \cdot \overline{\boldsymbol{\theta}})$, the "angle" between $\boldsymbol{\theta}$ and $\overline{\boldsymbol{\theta}}$;
- $\mathcal{E}_{\boldsymbol{v}} = ||\overline{\boldsymbol{v}} - \boldsymbol{v}||$;
- $\mathcal{E}_{\boldsymbol{\omega}} = ||\overline{\boldsymbol{\omega}} - \boldsymbol{\omega}||$.

Here $\mathcal{E}_{\boldsymbol{\theta}}$ is a measure of the accuracy of the estimates of $\boldsymbol{C} : \boldsymbol{W}$, and $\mathcal{E}_{\boldsymbol{v}}$ and $\mathcal{E}_{\boldsymbol{\omega}}$ are derived from $\boldsymbol{\theta}$, but are not uniformly sensitive to errors in each component of $\boldsymbol{\theta}$. Note that an estimate which is more accurate in terms of $\mathcal{E}_{\boldsymbol{\theta}}$ may fail to be more accurate in terms of $\mathcal{E}_{\boldsymbol{v}}$ and $\mathcal{E}_{\boldsymbol{\omega}}$.

## 5.3. Performance analysis for synthetic optical flow data

### 5.3.1. Comparing cost functions or minimisation techniques

We now compare the performance of several combinations of cost function and minimisation technique using synthetic optical flow data. In doing so, we use the following procedure:

1. Generate ideal optical flow data and a noise model as described in Section 4.1, with the standard deviation of each component of noise set to a random value between 0.01 and 1.0 pixels.
2. Generate one hundred sets of noisy optical flow data based on this ideal data and the noise model.
3. For each combination of cost function and minimisation scheme in question:
   (a) Compute $\overline{\boldsymbol{\theta}}$, $\overline{\boldsymbol{v}}$ and $\overline{\boldsymbol{\omega}}$ for each noisy optical flow field.
   (b) Determine the average value of $\mathcal{E}_{\boldsymbol{\theta}}$, $\mathcal{E}_{\boldsymbol{v}}$ and $\mathcal{E}_{\boldsymbol{\omega}}$ over the one hundred optical flow fields.

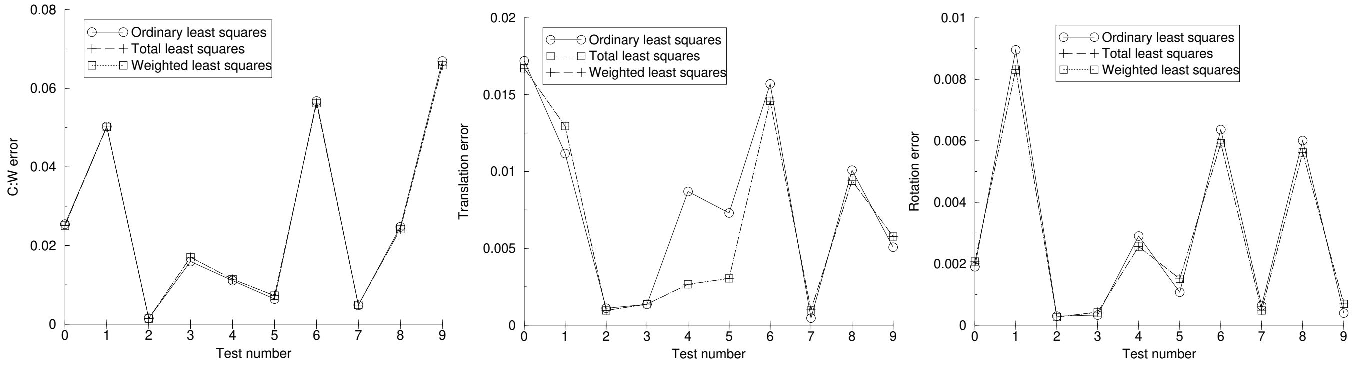


**Figure 2.** Comparison of cost functions given homogeneous isotropic noise.

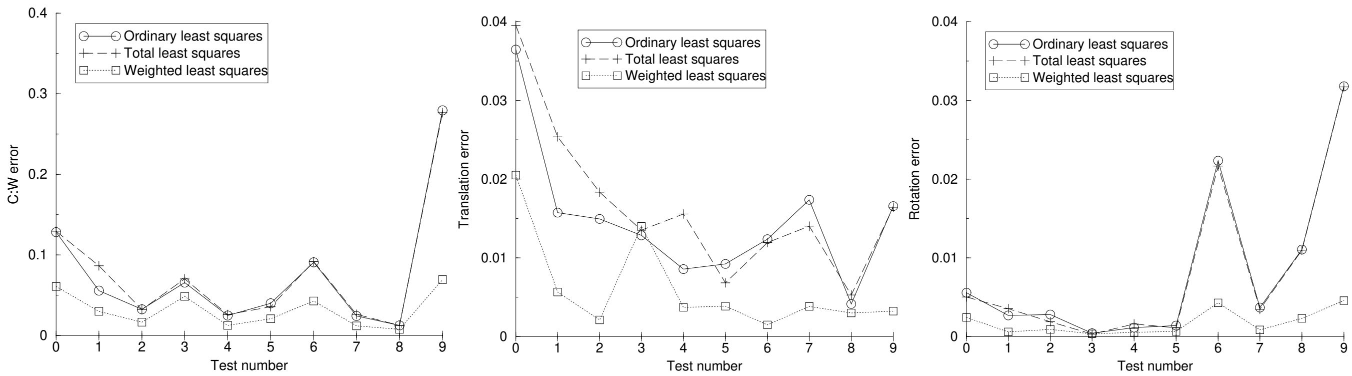


**Figure 3.** Comparison of cost functions given inhomogeneous anisotropic noise.

4. Repeat Steps 1, 2 and 3 for ten different sets of optical flow data and noise models.

From this procedure we obtain ten comparisons of the average performance of each combination of cost function and minimisation scheme for different ideal data and noise models.

Note that if we were engaged in real image tests, then we would in addition have to apply, as a first step, an appropriate technique for removing outliers.[7]

### 5.3.2. Comparison of cost functions

We first compare the performance of each of the three cost functions OLS, TLS and WLS defined in Section 3 when minimised using Sampson's Method.

Figure 2 shows a typical set of results obtained in the presence of noise which is constrained to be homogeneous and isotropic, while Figure 3 gives another set of results obtained in the presence of noise without this constraint. Although we only show results obtained using Sampson's Method, these results are similar to those obtained using other minimisation techniques. In particular, the following points apply regardless of the method used:

- In almost all cases, there is no significant difference between $\mathcal{E}_{\boldsymbol{\theta}}$ for the OLS function and the TLS function. This is due to the fact that the denominator of the TLS function varies very little between different elements, so that the TLS function is close to being a scalar multiple of the OLS function. There is a noticeable difference between $\mathcal{E}_{\boldsymbol{v}}$ for these cost functions in the presence of general noise, but because neither is consistently better we consider this to be due to the sensitivity of $\mathcal{E}_{\boldsymbol{v}}$ to small changes in the estimate of $\boldsymbol{C} : \boldsymbol{W}$.

- In the case of homogeneous isotropic noise, the WLS function reduces to the TLS function, as described in Subsection 3.3. Thus we would expect these functions to yield very similar results for such noise, and this is borne out by Figure 2.

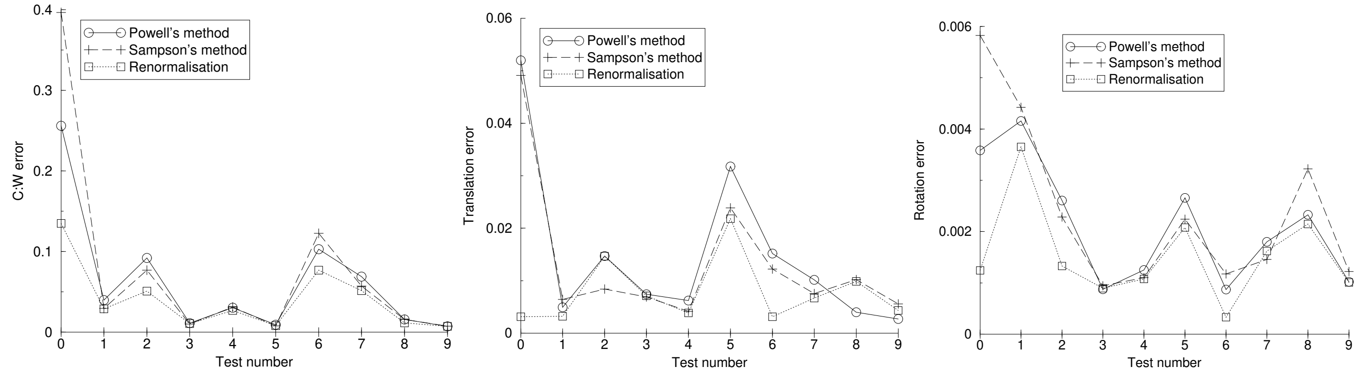


**Figure 4.** Comparison of three techniques for minimising the WLS cost function given inhomogeneous anisotropic noise.

- It is in the presence of inhomogeneous and anisotropic noise that the WLS method proves superior. It is clearly able to use the extra information about the noise to improve its accuracy in terms of all three error measures.

### 5.3.3. Comparison of Minimisation Techniques

Having ascertained that the WLS cost function is the most effective in the presence of general noise, we now determine the best available means of minimising it. Figure 4 gives a typical comparison of the errors produced by Sampson's method, Powell's method and renormalisation when minimising the WLS expression.

Renormalisation performs markedly better in terms of $\mathcal{E}_{\boldsymbol{\theta}}$, which implies that it generally estimates the components of $\boldsymbol{C}$ and $\boldsymbol{W}$ more accurately. Renormalisation also minimises $\mathcal{E}_{\boldsymbol{v}}$ or $\mathcal{E}_{\boldsymbol{\omega}}$ more effectively, although a little less convincingly, perhaps due to the extreme sensitivity of the estimates $\boldsymbol{v}$ and $\boldsymbol{\omega}$ to small errors in $\boldsymbol{C} : \boldsymbol{W}$.

## 6. CONCLUSION

The structure from motion problem of self-calibrating a moving camera from instantaneous optical flow was revisited. Central to this was the relationship between optical flow, camera motion and intrinsics specified by a differential epipolar equation and a cubic constraint. It was assumed that the observed data were subject to contamination, and a model of noise was duly adopted. It was further assumed that, for each piece of data, a covariance matrix was available describing the uncertainty of the datum.

Within this framework, three problem formulations were proposed, termed ordinary least squares, weighted least squares and total least squares. Each of these was evolved from an algebraic residual arising out of the differential epipolar equation. The cubic constraint did not enter any of these formulations. Informal tests have shown that a posterior corrective application of the constraint generates little discernible benefit. More work is needed to fully incorporate the cubic constraint into the estimation process.

Having obtained three least-squares cost functions, three minimising schemes were adopted as a means of obtaining desired solutions: Sampson's method, Powell's method and Kanatani's renormalisation. Synthetic tests revealed that there was little difference in performance between any of the combinations of cost function and minimising scheme unless the data were inhomogeneous and anisotropic in character. In this case, the best performer was the renormalisation technique associated with the weighted least squares cost function.

## REFERENCES


1. M. J. Brooks, W. Chojnacki, and L. Baumela, "Determining the egomotion of an uncalibrated camera from instantaneous optical flow," *Journal of the Optical Society of America A* **14**(10), pp. 2670–2677, 1997.
2. T. Viéville and O. D. Faugeras, "Motion analysis with a camera with unknown, and possibly varying intrinsic parameters," in *Proceedings of the Fifth International Conference on Computer Vision*, pp. 750–756, IEEE Computer Society Press, Los Alamitos, CA, (Cambridge, MA, June 1995), 1995.

3. N. Ohta and K. Kanatani, "Optimal structure-from-motion algorithm for optical flow," *IEICE Transactions on Information and Systems* **E78-D**, pp. 1559–1566, December 1995.
4. N. C. Gupta and L. N. Kanal, "3-D motion estimation from motion field," *Artificial Intelligence* **78**(1-2), pp. 45–86, 1995.
5. D. J. Heeger and A. D. Jepson, "Subspace methods for recovering rigid motion I: algorithm and implementation," *International Journal of Computer Vision* **7**(2), pp. 95–117, 1992.
6. P. A. Beardsley, A. Zisserman, and D. W. Murray, "Sequential updating of projective and affine structure from motion," *International Journal of Computer Vision* **23**(3), pp. 235–259, 1997.
7. M. J. Brooks, W. Chojnacki, L. Baumela, and A. van den Hengel, "Robust techniques for the estimation of structure for motion," in *Computer Vision—ECCV'98*, H. Burkhardt and B. Neumann, eds., vol. 1406 of *Lecture Notes in Computer Science*, pp. 281–295, Springer, Berlin, (Fifth European Conference on Computer Vision, Freiburg, Germany, June 2–6, 1998), 1998.
8. Q.-T. Luong, R. Deriche, O. D. Faugeras, and T. Papadopoulo, "On determining the fundamental matrix: analysis of different methods and experimental results," Tech. Rep. 1894, Institut National de Recherche en Informatique et en Automatique, Sophia Antipolis, France, Apr. 1993.
9. Q.-T. Luong and O. D. Faugeras, "The fundamental matrix: theory, algorithms, and stability analysis," *International Journal of Computer Vision* **17**(1), pp. 43–75, 1996.
10. Z. Zhang, "Parameter estimation techniques: a tutorial with application to conic fitting," *Image and Vision Computing* **15**(1), pp. 57–76, 1997.
11. P. H. S. Torr and D. W. Murray, "The development and comparison of robust methods for estimating the fundamental matrix," *International Journal of Computer Vision* **24**(3), pp. 271–300, 1997.
12. K. Kanatani, "3-D interpretation of optical flow by renormalization," *International Journal of Computer Vision* **11**(3), pp. 267–282, 1993.
13. K. Kanatani, *Statistical Optimization for Geometric Computation: Theory and Practice*, Elsevier, Amsterdam, 1996.
14. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, Cambridge University Press, Cambridge, 1995.
15. P. D. Sampson, "Fitting conics sections to 'very scattered' data: An iterative refinement of the Bookstein algorithm," *Computer Graphics and Image Processing* **18**(1), pp. 97–108, 1982.